Aalto University
School of Business

ADVANCED ANALYTICS SUCCESS FACTORS

A Case Study

Master's Thesis
Olli-Pekka Mattila
Aalto University School of Business
Information and Service Management
Autumn 2018

Aalto University
School of Business

Aalto University, P.O. BOX 11000, 00076 AALTO
www.aalto.fi

Abstract of master's thesis

Author Olli-Pekka Mattila

Title of thesis Advanced Analytics Success Factors – A Case Study

Degree Master of Science in Economics and Business Administration

Degree programme Information and Service Management

Thesis advisor(s) Juuso Liesiö and Eeva Vilkkumaa

Year of approval 2018 **Number of pages** 62 **Language** English

Abstract

Companies are increasingly taking into use advanced analytics solutions. Advanced analytics solutions are computer programs that analyze data, make predictions on the future, and give optimization-based recommendations on courses of action for achieving pre-determined business goals. Analytics solutions employ sophisticated statistical and mathematical models, and are often offered by third parties. Companies use analytics solutions to improve the efficiency of their operations.

This thesis studies whether the distinction between analytics and advanced analytics made in literature is well-founded. The second aim of this study is to find out, what contributes to an analytics initiative's success.

The study begins with a literature review synthesizing the findings of previous analytics research. The resulting synthesis identifies four distinct stages in an analytics project. They are acquiring data, transforming it into insights, communicating the insights, making business decisions, and finally implementing the decisions. Factors that contribute to each stage's success are identified.

The hypotheses that were developed in the theoretical part of the thesis are subsequently tested empirically using the single case study method and semi-structured interviews.

The case study confirms the findings of earlier research. Analytics can be viewed as a process with clearly identifiable stages. Specific measures can be taken to improve the success of each stage. The results obtained suggest that an analytics initiative should always be preceded by a thorough goal definition stage. This is a finding that earlier research has not emphasized sufficiently.

The study offers business executives a clear roadmap for managing analytics initiatives. It formulates clear action points and allocates parties the responsibility for executing them. The study also highlights some ordinary pitfalls preventing companies from fully benefitting from the results of analytics initiatives.

Finally, the study points out new interesting research opportunities in the intersection of information systems science and cognitive science. A key difficulty in using analytics effectively is that the reasoning behind the insights created by the solutions are often complex. Cognitive science could provide us tools for making the insights easier to digest. Lastly, the study highlights that process decoupling will eventually be applied to analytics initiatives. Future studies should research how the stages of an analytics initiative can be separated from each other, and outsourced to parties performing them the most effectively.

Keywords advanced analytics, analytics

Maisterintutkinnon tutkielman tiivistelmä

Tekijä Olli-Pekka Mattila

Työn nimi Advanced Analytics Success Factors – A Case Study

Tutkinto Kauppatieteiden maisteri

Koulutusohjelma Tieto- ja Palvelutalous

Työn ohjaaja(t) Juuso Liesiö ja Eeva Vilkkumaa

Hyväksymisvuosi 2018 **Sivumäärä** 62 **Kieli** Englanti

Tiivistelmä

Yritykset ottavat kiihtyvällä tahdilla käyttöön edistyksellisiä analytiikkaratkaisuja. Edistykselliset analytiikkaratkaisut ovat tietokoneohjelmia, jotka analysoivat liiketoimintaa koskevaa dataa, ennustavat datan perusteella tulevaisuuden kehityskulkuja ja antavat toimintasuosituksia yrityksen tavoitteiden saavuttamiseksi. Analytiikkaratkaisut perustuvat monimutkaiselle matemaattiselle ja tilastotieteelliselle mallinnukselle ja niitä hankitaan usein ulkopuolisilta järjestelmätoimittajilta. Yhtiöt käyttävät analytiikkaratkaisuja tehostaakseen toimintaansa.

Tässä opinnäytetyössä tutkitaan, voidaanko analytiikka ja edistyksellistä analytiikkaa pitää erillisinä ilmiöinä, kuten kirjallisuudessa on esitetty. Tutkielman toisena tutkimuskysymyksenä on selvittää, miten analytiikkahankkeiden menestystä voidaan edesauttaa.

Tutkimus alkaa kirjallisuuskatsauksella, jonka lopputuloksena aiemmat tutkimustulokset syntetisoidaan. Synteesin mukaan analytiikkahankkeessa on viisi osavaihetta. Osavaiheet ovat datan hankinta, datan analyysi, suositusten laadinta, liiketoimintapäätösten teko ja päätösten toimeenpano. Myös jokaisen osavaiheen onnistumiseen vaikuttavat tekijät määritellään.

Tutkimuksen kirjallisuuskatsauksen perusteella kehitettyä tulkintakehystä testataan empiirisesti tapaustutkimuksen välinein puolistrukturoiduin teemahaastatteluin.

Tapaustutkimus vahvistaa kirjallisuudessa esitettyjen johtopäätöksen paikkansapitävyyden. Analytiikkahanketta voidaan tarkastella prosessina, jossa on selvästi toisistaan eroavia vaiheita. Analytiikkahankkeiden menestystä voidaan parantaa konkreettisin vaihekohtaisin toimenpitein. Tutkimustulosten perusteella huolellisen tavoitteiden määrittelyvaiheen on aina edellettävä analytiikkahanketta. Aiemmassa kirjallisuudessa ei ole painotettu tämän löydöksen merkittävyyttä riittävästi.

Tutkimus tarjoaa liikkeenjohtajille konkreettisia neuvoja analytiikkahankkeiden johtamiseen. Tutkimuksessa määritellään selviä toimenpiteitä ja niiden toteuttamisesta vastaavat tahot. Opinnäytetyössä kartoitetaan myös tavallisia sudenkuoppia, jotka voivat estää yritystä hyödyntämästä täysimääräisesti analytiikkahankkeen tuloksia.

Lopuksi tutkielma nostaa esiin uusia mielenkiintoisia tutkimusmahdollisuuksia tietojärjestelmätieteen ja kognitiotieteen yhtymäkohdassa. Analytiikan tuottamien suositusten taustalla oleva päättely on usein monimutkaista. Kognitiotieteestä saattaisi löytyä keinoja monimutkaisen tiedon käyttämisen helpottamiseksi. Viimeisenä tutkimuksessa esitetään, että prosessien osien ulkoistamista tullaan väistämättä soveltamaan myös analytiikkahankkeissa. Tulevissa tutkimuksissa olisi hyvä selvittää, kuinka analytiikkahankkeiden vaiheet voidaan erottaa toisistaan ja ulkoistaa kolmansien osapuolten hoidettaviksi.

Avainsanat edistyksellinen analytiikka, analytiikka

Table of Contents

List of Figures

List of Tables

1 Introduction

"Analytics don't work at all. It's just some crap that people who were really smart made up to try to get in the game because they had no talent." (Charles Barkley, NBA hall of famer)

1.1 Research gap and research questions

This master's thesis studies advanced analytics (AA) and especially what affects the success in implementing advanced analytics solutions. Briefly put, advanced analytics solutions are computer programs, which analyze data, transform it into insights that predict the future operating environment, and recommend courses of action for reaching the company's goals. Advanced analytics solutions enable companies to allocate their resources more efficiently, and consequently improve their financial results.

The implementation of advanced analytics solutions on a wide scale is likely to change some of the traditional premises of conducting business. Business-decision making has traditionally been a realm of humans and taken place in the company. The automation of core business processes by the means of implementing advanced analytics solutions is likely to radically challenge the two main premises. The automation of processes is likely to reduce the sphere of human decision-makers' influence by moving parts of human decision-making to computers. Also, the automated processes are increasingly managed by external parties in their proprietary advanced analytics solutions. Despite of the scale of potential change and the significant financial interests in play, a surprisingly small number of studies has previously dealt with the topic.

Stockmann Herkku, a Finnish grocery store chain owned by the retail conglomerate S-Group serves as an excellent example of this trend. The company has recently licensed a supply chain management tool used to make automated replenishment orders and demand estimates from Relex, the leading Finnish retail analytics provider (Relex 2018). The company has also licensed another analytics solution for pricing its grocery products in promotion and a third tool for allocating their marketing investments from a smaller provider of analytics solutions (Sellforte 2018).

When core business processes are increasingly managed with advanced analytics tools, sometimes provided by third parties, understanding what contributes to their implementation success becomes critical. If companies do understand how to implement best-of-breed advanced analytics solutions successfully, they are likely to gain additional competitive advantage by operating more efficiently and meeting the market's wishes more accurately.

The research question of what affects advanced analytics solutions' implementation success has not yet been studied in a theoretically and empirically rigorous manner. The shortcomings of literature can be attributed to at least three specific reasons.

The existing articles have approached analytics either from a too holistic or piecemeal fashion or alternatively lacked convincing empirical support. Some notable articles view analytics from a broad perspective and have mainly attempted to provide starting points for further studies of analytics. (see e.g. Vidgen *et al.* 2017, Holsapple *et al.* 2014). On the other hand, in some articles the views have been rather narrow, only analyzing analytics from one stakeholder group's view, like Kowalczyk and Buxmann's article that focused on analysts' views (Kowalczyk and Buxmann 2015). Such approaches are not likely to be generalizable. Lastly, surprisingly few articles have explicitly addressed advanced analytics because most often the focal point of research has been on general analytics, like in Seddon *et al.*'s article (Seddon *et al.* 2001).

The scarcity of source material on the subject matter of this thesis may also be partly explained by the fact that analytics as a wider phenomenon has not been studied much. Amongst others, Mortenson et al. have pointed out that a surprisingly low number of analytics-related research has been published in the operations research/management science journals (Mortenson et al. 2014). It should be mentioned that the number of articles covering analytics appears to be rising, at least modestly. For example, during the writing process of this thesis covering the time period between June 2017 to May 2018 new noteworthy articles - like Vidgen's *et al.*'s study (Vidgen *et al.* 2017) - were published in leading OR journals.

Further, Ranyard et al. have criticized the research practices of the operations science community. According to the authors, there exists a gap between theory and practice in operations management/management science. Allegedly, academics have a preference for tackling difficult theoretical questions having few practical applications, whereas easier but more practically oriented questions are left unaddressed. (Ranyard et al. 2015) In comparison to the traditional hard quantitative operations science themes, the qualitative research question of this thesis appears to be a relatively simple one. Anyhow, its practical implications are important.

This thesis aims to contribute to lacking research of advanced analytics and to bridge the theory-practice gap by addressing a topic that is increasingly important to the practitioner community. Practical insights are acutely needed on how advanced analytics solutions can be successfully implemented.

This study aims to find answers find to the following two practical questions (i) are there any intrinsic characteristics of advanced analytics solutions that separate them from the other fields of analytics; (ii) which factors affect advanced analytics solutions' implementation success.

1.2 Methods

This thesis applies the single case study method with some additional elements adopted from the systematic combining research tradition (Dubois and Gadde 2002). Case study method was chosen because prior research on the subject matter of this thesis is relatively scarce. The case company is a small Finnish limited liability company operating in the field of advanced analytics. The company provides its proprietary advanced analytics solution to large domestic and international retail customers. Semi-structured interviews were the source of information of this thesis. Three employees of the case company were interviewed. Also, to triangulate the findings, a single semi-structured interview conducted with a person with a similar professional and educational background, who works in a different environment, a pensions insurance fund. The interview transcripts were thematically coded and analyzed by comparing their contents to the theoretical synthesis made in the literature review part of this thesis.

1.3 Structure of the thesis

The first chapter of this thesis explains the rationale for conducting this study and sets outs the general goals and specific research questions of this thesis. The second chapter contains the literature review part familiarizing the reader with the field of analytics and advanced analytics. First, the essential characteristics of analytics as a research topic are identified on the basis of academic literature and the difference between analytics and advanced analytics is discussed. After that, I will describe the factors facilitating – and correspondingly hindering – the adoption of analytics in business organizations. Finally, the insights of the review are synthesized into an unified advanced analytics implementation success framework.

In the third chapter, the reader is introduced to the methodological background of this thesis – single case study method – that has been used to conduct this study. The research procedure, including a description of research data sources, data collection methods, and the method of analysis are set out in the third chapter.

The fourth chapter analyzes and discusses the findings of the study. The last, i.e. the fifth chapter contains the conclusions of this thesis and discusses the general relevance of the findings, assesses their limitations and proposes new topics for further research.

2 LITERATURE REVIEW

2.1 General remarks on analytics

2.1.1 The concept of (business) analytics

This master's thesis studies advanced analytics. To conduct meaningful research, one must thoroughly understand what the concept advanced analytics means and what not, and how the concept relates to the other branches of analytics.

This is especially important because academic business literature – not to even mention everyday business jargon – seems to contain a variety of almost identical terms relating to analytics, whose definitions are often overlapping or unclear. Such terms include, just to give a few examples, analytics (Davenport 2007), business analytics (Evans 2012, Vidgen 2017, Holsapple *et al.* 2014, Seddon *et al.* 2017), and business intelligence and analytics (Kowalczyk and Buxmann 2015, Chen *et al.* 2015). Some authors, like Bayrak uses business analytics, business intelligence and big data synonymously (Bayrak 2015), whilst Holsapple *et al.* deem that analytics and business analytics refer to the same subject (Holsapple *et al.* 2014). A prime example of terminological inconsistency found during the writing process of this thesis was when one author (Bose 2009) managed to define advanced analytics in three different ways on four consecutive pages in the same article. The problem of terminological inconsistency plaguing the field of analytics has also been recognized in literature (Mortenson *et al.* 2015).

Besides the terminological issues, the methodological grounds of analytics are subject to an ongoing discussion. In addition to the traditional quantitative methods, analytics is often claimed to also encompass the so-called SoftOR, i.e. problem structuring methods, which are qualitative techniques used to understand and solve problems having multiple stakeholders and objectives (Ranyard *et al.* 2015, Holsapple *et al 2014)*. The relationship between the SoftOR methods and the subject of this thesis warrants clarification.

In order to clearly establish the borders of the subject-matter of this study, the reader will be next introduced to the field of (business) analytics, which will ultimately enable us to define what advanced analytics encompasses. In the second chapter, the success factors of analytics projects are presented to the reader. The following analysis starts from definitional considerations, followed by a review of previous analytics-related models and frameworks, and finally proceeds to the detailed analysis of literature.

2.1.2 Analytics is a subset of business intelligence and analytics (BI&A)

The majority of authors define the difference between business analytics and intelligence (BI&A) and analytics as follows: (business) analytics is concerned with the process of transforming data into insights to make decisions, whereas the field of business intelligence encompasses analytics and additionally the study of technologies, systems, and applications used to analyze business data to make better business decisions (Chen *et al.* 2012, Kowalczyk and Buxman 2015, Seddon *et al*. 2016).

The topic of this thesis, advanced analytics, is inseparably connected to the underlying technologies, because AA solutions run on computers. For this reason, this thesis could be regarded to belong either to analytics research, or business intelligence and analytics research.

However, this thesis will not study the AA solutions' technological infrastructure, which is more suitably addressed in engineering-oriented disciplines and computer science. In this thesis, the technological aspects of the phenomenon are only analyzed from a descriptive point of view, aiming only to give the reader a working knowledge of the technological infrastructure. As a consequence, this thesis falls under the scope of analytics research.

2.1.3 Analytics means transforming data into insights for making better decisions

INFORMS, the world's leading community of analytics and operations research professionals defines analytics as "*the scientific process for transforming data into insights for making better decisions*" (INFORMS 2017). Liberatore and Luo have given analytics the following meaning "*a process of transforming data into actions through analysis and insights in the context of organizational decision making and problem solving*" (Liberatore and Luo 2010). Davenport and Harris define analytics in the following way: "*the extensive use of data, statistical analysis, explanatory and predictive models, and fact-based* *management to drive decisions and actions*." (Davenport and Harris 2007) According to Holsapple *et al.* "*business analytics is concerned with evidence-based problem recognition and solving that happen within the context of business situations*." Many authors, like Holsapple *et al.,* use analytics as a shorthand expression for the term business analytics. (Holsapple *et al.* 2014*)*

2.2 Previous theories on analytics

In literature, several frameworks and models have been developed for understanding analytics. Practically every article published in leading academic journals has presented its readers a new framework, either based on the authors' own work or building on earlier frameworks. To the knowledge of the writer of this thesis, none of the published frameworks has subsequently been adopted in academic or practitioner communities, except for as a basis for further framework development. This sub-chapter contains very brief descriptions of the analytics models and frameworks found during the research process.

The numerous analytics models and frameworks found in literature include, to name but a few, Seddon *et al.*'s BASM (Seddon *et al.* 2016), Holsapple *et al.*'s BAF framework (Holsapple *et al.* 2014), Kowalczyk and Buxmann's model of the theory of ambidexterity in decision support (Kowalczyk and Buxmann 2015), Davenport and Harris's DELTA framework as reported in Seddon *et al.'s* article (see Seddon *et al.* 2016), Davenport *et al.*'s model for building an analytic capability (Davenport et al. *2001*), and finally Liberatore and Luo's process view of analytics (Liberatore and Luo 2010). In addition, many articles study analytics projects' success factors without setting out formal models or frameworks.

Seddon *et* al.'s abbreviation BASM stands for business analytics success model. The model is a synthesis of earlier research building on 16 older models, including for example, Seddon's own works and Davenport and Harris's DELTA framework, DELTA referring to data quality, enterprise-wide integration, leadership, well-chosen targets and analytic people. To be precise, BASM is not a single model but actually contains two sub-models. According to the authors, the models be used to analyze the success factors of a single analytics projects and also to identify the factors to improve an organization's long term analytics success. The BASM's sub-models were tested empirically by validating them against analytics solutions providers' marketing materials (Seddon *et al*. 2016) Seddon *et al.*'s article contains the most extensive listing of analytics' projects' success factors in literature that was reviewed. Therefore, the article is included in the analysis. However, due to the questionable objectivity of the underlying empirical evidence – supplier marketing material, which may intentionally twist reality in favor of the suppliers – the authors' findings must be treated with due care and reservation.

Holsapple *et al.*'s BAF framework contains six perspectives to business analytics. Holsapple *et al.'s* six perspectives were synthesized as a result of analyzing 18 sources containing differing definitions of business analytics. Thus, the framework is a result of purely conceptual contemplations having no empirical support. The possible perspectives to analytics are to view it as a movement, decisional paradigm, specific activities, practices and technologies, transforming process, or a capability set. According to authors, the BAF framework can be used by practitioners in the planning and evaluation of analytics initiatives. The authors further argue that the framework provides the academic community with a structured research agenda for further studies of the topic. (Holsapple *et al. 2014)* Due to its conceptual breadth, depth, and especially SoftOR-positive orientation, the article and the BAF model serve as particularly good yardsticks in the following development and clarification for the concept of analytics.

Kowalczyk and Buxmann's model of the theory of ambidexterity in decision support identifies six tensions, which may adversely influence the success of analytics projects. The article develops tools and tactics for mitigating these tensions, thus increasing the rationality and quality of the organizational decision-making process. The authors' findings are based on multiple case study interviews of analysts, which lends credibility to the authors' viewpoints. (Kowalczyk and Buxmann 2015) Parts of the model will be included in this thesis, since the article provides the most extensive empirically validated description of the decision stage of analytics. Due to its narrow focus on analysts' role in analytics initiatives, the article may not provide a comprehensive framework for understanding analytics initiatives holistically.

Davenport *et al.*'s model for building an analytic capability analyzes the factors that affect the success of analytics initiatives on three levels. On the first level of analysis, the contextual factors that shape the context, in which analytics projects are executed are addressed. These factors include strategy, employee skills, culture and available data and technology. Davenport *et al.* argue that the factors are the prerequisites of analytics success. The second level of analysis focuses on the transformational phase of the process. It identifies factors that affect the process of turning data into insights and making decisions on the basis of the results of the analysis. On the third and last level of analysis, the authors describe the factors that determine the success of implementation. The success of implementation can be assessed on the basis of financial, behavioral and processual outcomes. (Davenport et al. *2001*) Davenport *et al.*'s insights were validated by interviewing a large number of subject matter experts, which lends credibility to their conclusions.

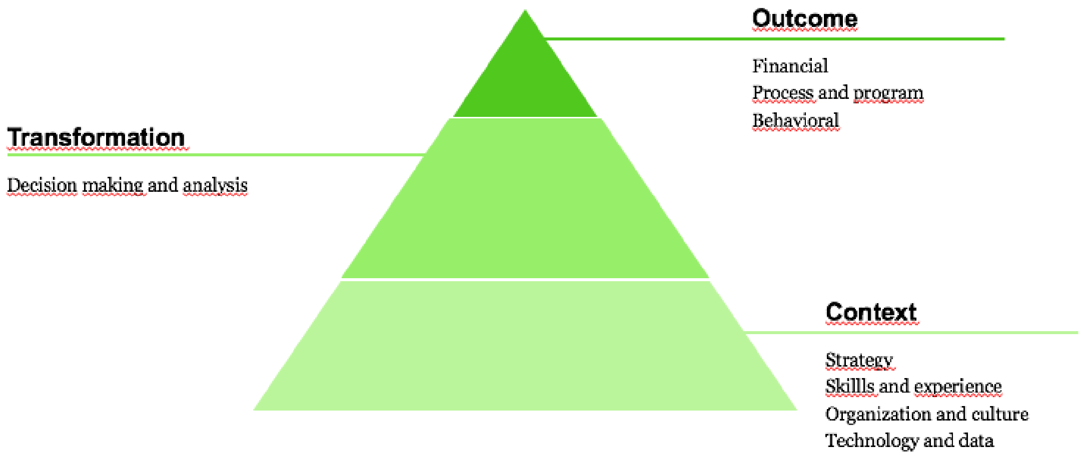

Figure 1: A Model for Building an Analytic Capability (Davenport *et al*. 2001)

Lastly, Liberatore and Luo have developed a process view to analytics. According to the authors' process view, analytics initiatives have four stages. The first stage of analytics is obtaining the necessary data and preprocessing it for analysis. The data is analyzed in the second stage of the process. Insights emerging from the analysis aid in making sense of the operating environment. Finally, business decisions are made and actions taken on the basis of the insights. In addition, increased amount of data, analytic people, software, and organizational process orientation in operations are argued to drive analytics adoption and success. (Liberatore and Luo 2010) The authors viewpoints are purely conceptual but yet provide an elegant and holistic way to view analytics.

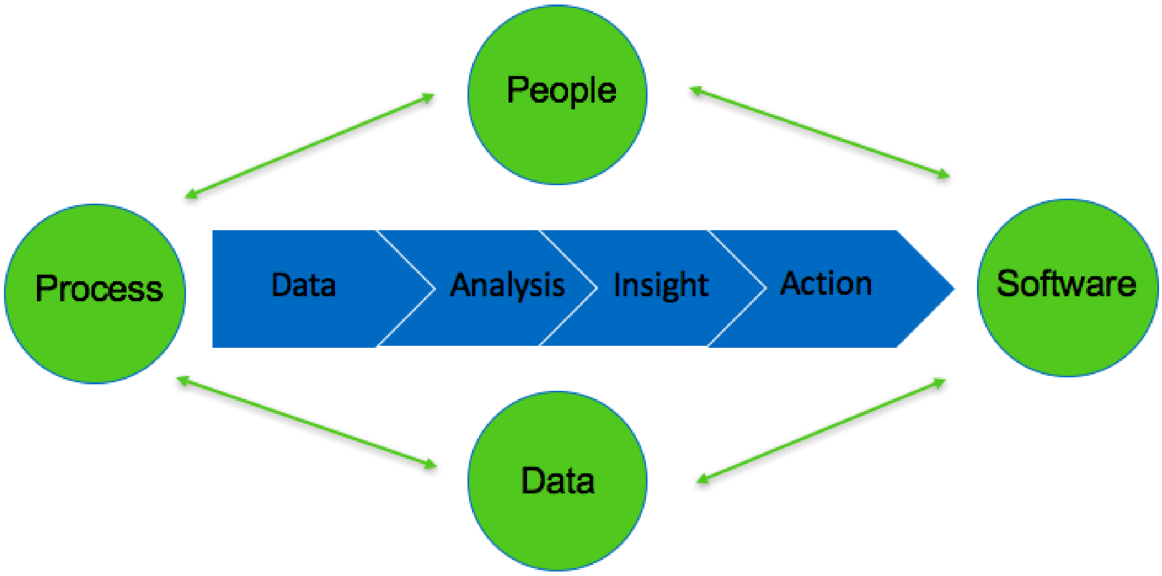

Figure 2: The stages of analytics and its driving forces of analytics, adapted from Liberatore and Luo 2010.

2.3 Preliminary framework adaptation for the thesis

In this sub-chapter, we will briefly compare the key characteristics of the models and frameworks described in the above sub-chapter. The comparison is made in order to identify similarities between the models and frameworks. Once we have identified the common characteristics of the models and frameworks, the insights will allow us to formulate rough understanding of the stages of analytics and the factors contributing to the success of the analytics process. The sketched thematic skeleton will then serve us as a thematic device in the following detailed analysis of analytics.

In absence of a unifying preliminary framework, setting the findings of the subsequent extensive literature review in their proper contextual places could prove to be difficult. In the following chapters of the literature review, the thematic skeleton will be fleshed out with insights originating from the other frameworks and articles.

Many authors regard analytics as a process with clearly identifiable stages, most often three or four. Seddon *et al.* appear to argue that in the first stage of analytics organizations current analytical resources, such as high quality data and people are used first to produce insights (Seddon *et al.* 2016). Liberatore and Luo propose that analytics begins with data processing (Liberatore and Luo 2010) whereas Davenport *et al*. claim data processing to be an enabling factor of the process (Davenport *et al*. 2001). Thus, obtaining and pre-processing of data can be considered either as a stage or an enabling factor of analytics. In the preliminary framework construction, we initially deem data processing as the first stage of analytics, but will revert to this issue later on.

In the second stage of the process, the data is analyzed. This is a common feature of both Davenport *et al*.'s and Liberatore and Luo's models, as indicated in figures 1 and 2 above (Davenport *et al*. 2001, Liberatore and Luo 2010). In Seddon *et al*.'s model, people engage in "problem solving" by using enabling business intelligence technology, which arguably has the same meaning than what analysis has (Seddon *et al*. 2016). Thus, it looks like the authors' views are convergent on the matter, and therefore we define analysis as the second stage of analytics in our preliminary framework.

Liberatore and Luo and Seddon *et al.* identify insight generation as its own processual stage following the analysis step, whereas Davenport *et al.* do not regard it as a stage. Instead, Davenport *et al.* set in the article out that data is turned into useful insights and knowledge during the analysis stage. (Liberatore and Luo 2010, Seddon *et al*. 2016, Davenport *et al*. 2001). In the preliminary framework construction, we will regard insight generation as its own processual stage, even despite its relation to the analysis stage is yet unclear, and solve the matter later.

Lastly, all three author teams concur that business decisions are made on the basis of insights (Liberatore and Luo 2010, Seddon *et al*. 2016, Davenport *et al*. 2001). It should be noted that the views diverge on whether implementation should be regarded as its own processual stage being separate from managerial decision-making. Liberatore and Luo, unlike Davenport *et al.* and Seddon *et al*. do not seem to regard the implementation of the decisions as a stage. They take a more managerial perspective to decision-making, and seem to argue that it suffices that a managerial decision is made, whereas Davenport *et al.* emphasize the role of the employees in implementing the decision, while Seddon *et al.* highlight the importance of implementation without addressing who is in charge of the implementation. (Liberatore and Luo 2010, Seddon *et al*. 2016, Davenport *et al*. 2001) We will initially regard the decision-making and implementation as one stage, but will later resolve the question due to the points raised by the other authors.

Regarding the enabling factors, most authors share the view that data and technology are prerequisites of analytics capabilities. Liberatore and Luo regard that software and data are key drivers of analytics (Liberatore and Luo 2010). In Davenport *et al.*'s terminology technology and data are prerequisites of analytics, technology encompassing both hardware and software (Davenport *et al.* 2001). Both above-mentioned sources are in line with Holsapple *et al.*'s BAF setting out that analytics can be viewed from the perspective of practices and technologies (Holsapple *et al.* 2014). Data's importance has also been acknowledged by Seddon *et al.* (Seddon *et al.* 2016). For these reasons, our preliminary conception is that technology contributes to analytics success.

Liberatore and Luo name people as one of the driving forces of analytics, whilst Davenport *et al.* have labeled the same thing skills and experience, which are both attributable to people (Liberatore and Luo 2010, Davenport *et al.* 2001). Holsapple *et al.* do not regard that people contribute directly to analytics success but their impact is indirectly acknowledged when viewing analytics from the capability set perspective, which acknowledges that the capabilities of the organization and its individual operatives impact analytics success (Holsapple *et al.* 2014). Analytic people are also named as a key success factor by Seddon *et al.* (Seddon *et al.* 2016).

Third, Liberatore and Luo set out that process-orientation, i.e. measuring and focusing on the organization's value-adding activities supports analytics initiatives (Liberatore and Luo 2010). In Davenport *et al.*'s terminology strategy, which refers to identifying key processes and defining the key decisions regarding the processes contributes to analytics success (Davenport *et al.* 2001). Seddon *et al.* argue that the extent of embedding evidence-based decision making into an organization's processes affects analytics success (Seddon *et al.* 2016). Therefore, it seems clear that the extent to which the organization's members view the business as a process with its own metrics and controllable variables affects the success of analytics initiatives.

Fourth, many of the above models and frameworks set out that cultural factors influence analytics. The view is shared amongst other by Davenport *et al.*, Seddon *et al.* and

Holsapple *et al.* (Davenport *et al.* 2001, Seddon *et al.* 2016, Holsapple *et al.* 2014). The cultural factors will not be studied in this thesis. The reason for this limitation is that this study is concerned with one-off analytics initiatives and cultural changes may require a longer time-period to take place.

To conclude this sub-chapter, the main sources share very similar views regarding the key stages of analytics and the factors that contribute to the success of analytics initiatives. Whilst there undeniably are several ways to divide analytics into separate stages and group the factors influencing analytics under concepts of differing scopes, this thesis initially adopts a four-staged process-based view to analytics.

The stages of analytics are assumed to be data processing, analysis, insights, and decision-making / action-taking. Further, it is assumed that three key factors, technology, people, and process-orientation contribute to the success of analytics initiatives, cultural factors being explicitly disregarded. Next, we will begin defining the contents of the processual stages in detail and also attempt to gain a more granular picture into the key analytics success factors.

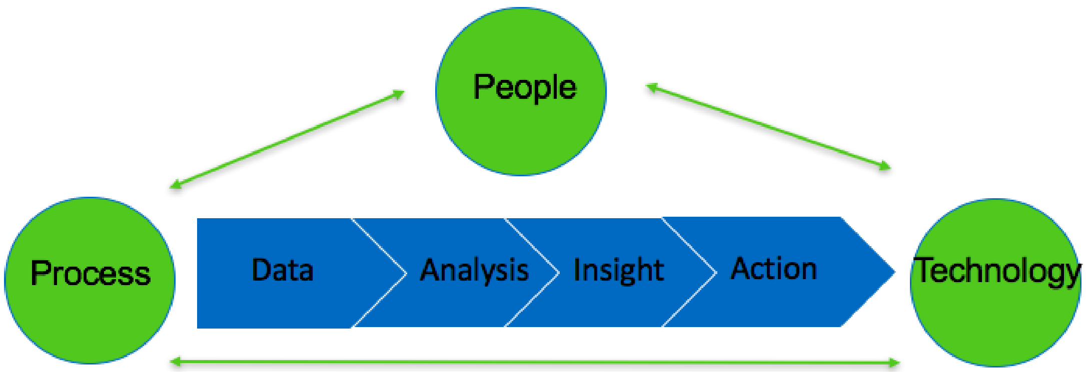

Figure 3: The process model of analytics and success factors

2.4 Analytics uses data as its input material

2.4.1 What is data?

According to our preliminary framework, analytics uses data as its input material. In the ordinary usage of language, the concept of data refers to electrically stored pieces of information representing a specific state of the world. The availability of good quality data has been consistently found to be a prerequisite of successful implementation of analytics projects (Seddon *et al*. 2016, Kowalczyk and Buxmann 2015). However, in some authors' view data availability is not the most important driver of analytics success (Lavalle *et al*. 2011).

Holsapple *et al.* have suggested a small modification to the definition of analytics According to said authors, instead of the term data, evidence is better suited to be used in the definition of analytics to describe the input material of the analytics process. The authors point out that hard facts or data are not always available in decision-making situations. As a consequence, the evidentiary enhancement would accommodate better the fact that analytics may also deal with justified estimates, unbiased observations and credible explanations and the like instead of only data or hard facts. In other words, the terminological improvement would allow analytics to draw on inputs which are uncertain. (Holsapple *et al.* 2014) Seddon *et al.* have also expressed a similar SoftOR-oriented view and characterized as analytics is "*intendedly rational problem solving*". Despite of all the available computing power, an analyst must engage in creative work to recognize and define the problem set, select the appropriate models and to finally assess the accuracy of the results. (Seddon *et al*. 2016)

Based on Holsapple *et al.*'s and Seddon *et al.*'s arguments, the question of analytics' input material does not seem to be as straightforward as what is perhaps suggested by for example Liberatore and Luo. Analytics appears to also encompass elements of subjective decision-making and uncertain inputs. However, if we are to accept the SoftOR-oriented views on their face value, we may run into other problems. If practically all business-related decision-making and problem solving are defined as analytics, that would inevitably dilute the differences between analytics and other forms of decision-making.

This thesis attempts to strike a balance between the traditional and SoftOR-oriented schools of analytics research. The modifications to the definition of analytics suggested by Holsapple *et al.* are not included in the definition of analytics of this thesis. As stated above, the first leg of the definition of analytics is that it deals with data, which is its input material. Holsapple *et al.*'s concerns do not necessitate tweaking that part of the definition.

Despite of this, the points of view expressed in the SoftOR-oriented literature cannot be overlooked. The process of analytics also contains subjective parts. In this thesis, the subjective parts are deemed to belong to the stage of the analytics process, where data is *transformed* into insights. This issue will be dealt with in chapter 2.5.

Despite of having now defined that analytics uses data as its input material, a broader understanding of the concept of data than the one contained in the above sources is still needed.

2.4.2 In this thesis data refers only to structured data

Data can be either structured or unstructured data. Structured data is labeled, i.e. the data relates to a predefined property, and it can be processed a computer. Semi-structured and unstructured data do not contain such labeling and are thus more difficult to process. (Baars and Kemper 2008, Gandomi and Haider 2015) Structured data is often stored in relational database management systems (Chen *et al.* 2010). Unstructured data, which includes e-mails, web pages and various social-media content (Baars and Kemper 2008, Apte *et al.* 2003) is not stored in a similar easy-to-access manner. Baars and Kemper further suggest that the technological requirements for processing structured and unstructured data differ from each other (Baars and Kemper 2008).

Another important conceptual distinction with respect to data relates to the differences between data and so-called big data. If data is particularly voluminous, highly variable by nature, and created very quickly in relation to what the existing data processing equipment capabilities allow, it is normally referred to as big data. The processing of big data is subject to its own special challenges. To give some examples, Gandomi and Haider have pointed out that big data is more heterogenous and noisy than ordinary data (Gandomi and Haider 2015). Gandomi and Haider as well as Mortenson *et al.* also note that big data inevitably contains more statistically significant spurious correlations than ordinary data does. For these reasons, all of the above authors share the view that new statistical techniques are needed to make predictive models and other OR/MS techniques operative in the big data environment. (Gandomi and Haider 2015, Mortenson *et al* 2015) In addition to methodological challenges, computational efficiency may also set limits for the use of predictive techniques in the big data environment (Gandomi and Haider 2015).

Due to the fact that the technical choices and statistical techniques applied in a big data environment differ from the ones employed in the “small” data environment, they will be carved out of the scope of this study. In this master’s thesis, the term data refers only to structured data that can be processed with state of the art data processing equipment and established quantitative techniques.

2.5 Analytics transforms data into insights

The second leg of the above-mentioned definitions of analytics and the second stage of our preliminary framework is that the data serving as the input of the process is *transformed* into something of greater value, *insights*. According to INFORMS’s definition and Liberatore and Luo’s views on analytics (INFORMS 2018, Liberatore and Luo 2010), the data explicitly goes through a transformative process. In their seminal book, Davenport and Harris acknowledge that technologies and processes use data to understand and analyze business performance (Davenport and Harris 2007), Thus, Davenport and Harris, Liberatore and Luo, and INFORMS all draw a clear conceptual distinction between the input material of the process i.e. data and the following transformative process.

However, if we attempt to understand the transformative process only by looking at the above definitions and the preliminary framework, the transformation stage still remains static and black-box-like. Data goes in the transformative process and emerges out of it as insights used to make to business decisions, but the steps preceding and taking place during the transformation process remain invisible and undefined. Thus, getting a better view of the specific contents of the transformative process is required for understanding analytics.

According to several sources, analysts are often unable to define the most suitable analytic methods in the beginning of the analytics process. Experimentation and iteration are regarded as necessary elements of developing the right approach to solve the problem at hand (Kowalczyk and Buxmann 2015, Viaene and Van der Bunder 2011, Davenport *et al.* 2001, Liberatore and Luo 2010). Similarly, Seddon *et al.* have suggested that analytics is essentially problem solving and highlight the importance of the creative human effort when problems are defined, formal solutions devised, and results analyzed and interpreted. (Seddon *et al.* 2016)

This thesis recognizes the importance of subjective and experimental human effort, when analytics solutions are designed, implemented, and their preliminary results assessed. In the view of the author of this thesis, the first part of the transformative process of analytics should also be regarded to encompass defining the problem, devising possible solutions to the problems, and finally, once the solution has been established, assessing the results critically. Also, the transformational process may not proceed in a linear way, but the problem-solving methods may in some instances have to be adapted in the light of later findings.

One further way to gain insights into the contents of the transformational process is take a look at the outputs of the process, which define what kind of methods need to be employed in the process (Holsapple *et al.* 2014). Analytics can be used to produce either descriptive, predictive, or prescriptive insights.

Descriptive analytics is concerned with questions such as what happened or what is happening. Descriptive analytics is normally understood to include standard business reporting and its aim is to identify business opportunities and problems. (Delen and Demirkan 2015) Said opportunities may be, for example, analyzing sales and classifying customers into different segments (Evans 2012).

Predictive analytics uses mathematical techniques, such as time series analysis, to uncover recurring patterns from the data to answer the question what will happen next. The aim of predictive analytics is to get a view on which factors affect the future courses of events and predict the future. (Delen and Demirkan 2015, Evans 2012) Predictive analytics also helps in assessing the probabilities of possible courses of events, which helps, for example in credit risk assessment (Evans 2012).

Finally, prescriptive analytics adds a normative element on the top of the predictive analytics layer. The normative element means that prescriptive analytics also provides recommendations concerning the measures that should be taken to reach an intended outcome taking into account the defined objectives and constraints of the scenario. (Delen and Demirkan 2015) In a business environment, predictive analytics is used to improve performance by either maximizing or minimizing some objective function (Evans 2012). Prescriptive modeling employs techniques like simulation, optimization, and multi-criteria decision modeling. (Delen and Demirkan 2015)

In the context of this study, I will refer to both predictive and prescriptive analytics with the term advanced analytics (AA). The term advanced analytics has been previously used before both in academic literature and business context. It should be noted that the meanings assigned to the term are not yet established but have varied, sometimes even within one article.

For example, Bose gave the term three differing meanings in the space of four consecutive pages. One the first page of the article, the term was used synonymously with predictive analytics (p. 155). Bose also defines advanced analytics tautologically as "*various advanced analytics techniques [--] used in combination with one another to gain information, analyze that information, and predict outcomes of that information*" (p. 156) and finally as "*a suite or cluster of analytical applications that helps measure, predict, and optimize organizational performance and customer relationships*" (Bose 2009, p. 155, 156, and 158). Such inconsistent use of terminology can be criticized because it only causes unnecessary confusion.

Other examples representing more practically oriented and established uses of the term advanced analytics can be found, for example, from Barton's and Court's article on advanced analytics, which appeared in Harvard Business Review in October 2012. Barton and Court do not explicitly define advanced analytics, but on the basis of the article that recommends the use role of predictive modeling and optimization in business it seems clear that they regard predictive models containing optimization elements as advanced analytics. (Barton and Court 2012). Barton's and Courts's definition appearing between the lines is similar to that used by Gartner, a consultancy, defining AA as (semi)automated examination of data using sophisticated techniques to make predictions and generate recommendations (Gartner 2018). Thus, the definition used in this thesis is consistent with the one found in literature and used in the relevant industry.

In addition to established industry terminology, combining both predictive and prescriptive analytics under one concept is arguably reasonable because prescriptive analytics applications have been argued to utilize insights generated by a predictive analytics models. Thus, the two modeling techniques seem to be inseparably tied to each other.

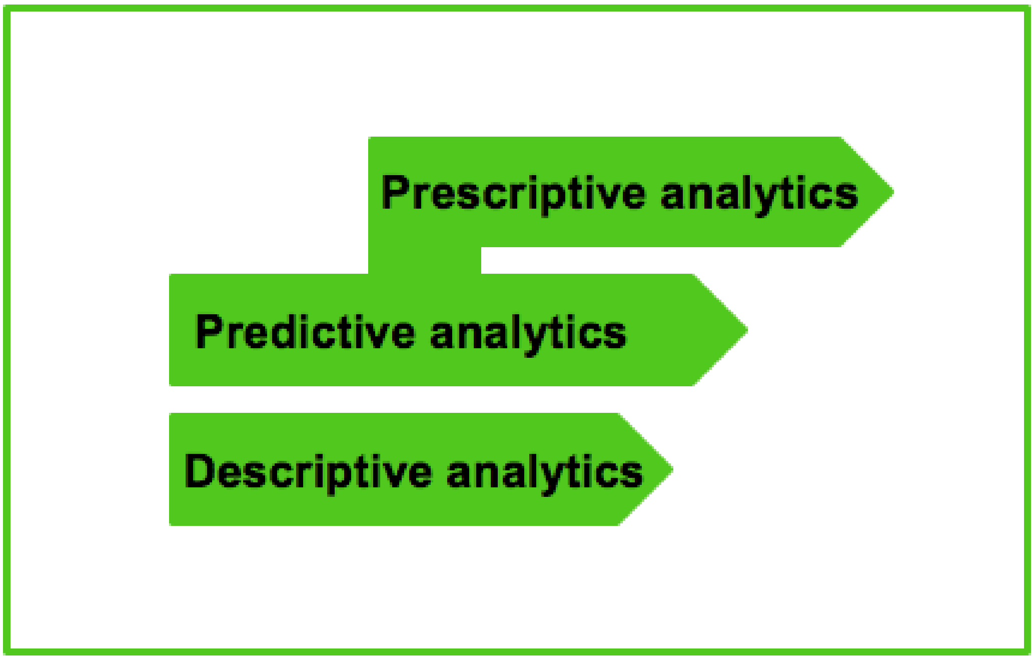

Figure 4: The three types of analytics

2.6 The transformational process produces insights

Third, our preliminary framework postulated that the transformational process produces *insights* as its output, which are used for an instrumental purpose of decision-making and/or action-taking in a business context in the next processual stage.

The majority of the key sources of this thesis share this view. For example, in INFORMS's view data is transformed into insights to make better decisions (INFORMS 2017) and Liberatore and Luo argue that analytics aims to transform data into actions through analysis and insights (Liberatore and Luo 2010).

The definition of insight has not received much direct attention in the analytics literature. This might be explained by the fact that the analysis of the nature of human insight normally falls under the scope of natural and behavioral sciences such as cognitive science, psychology, medical studies, and engineering. As of this date, human cognition's underlying physical mechanisms are not fully understood even in these fields. For this reason, in this thesis we will approach the matter in a pragmatic manner. We aim to understand what an insight is only to the extent it enables us to state something meaningful about AA implementation success.

Seddon *et al*. argue that insight refers to understanding, that is seeing the relationships between things clearly. The authors also set forth that the values of insights can be compared, some insights being more valuable than others (Seddon *et al*. 2016). Similarly, according to Liberatore and Luo, insights allow decision-makers to compare key performance indicators on multiple dimensions and identify root causes of problems (Liberatore and Luo 2010). However, the authors do not explain what makes some insights more valuable than other. Luckily, this question has been addressed in other analytics authors' articles.

Viaene and Van der Bunder have argued that it is of utmost importance to explain the thinking behind a decision to managers. This includes disclosing the model's assumptions. Otherwise the recommendations may not be acted upon. (Viaene and Van der Bunder 2011) Kowalczyk and Buxmann have stated that the quality of analytic advice, i.e. insights, can be signaled to decision-makers by using comprehensible models, demonstrating problem understanding, clarifying model assumptions, responding to decision makers' intuitions and using domain experts to communicate the results (Kowalczyk and Buxmann 2015). Davenport *et al.* emphasize the importance of appropriately packaging the outputs of analytics process and including a suitable proportion of text, numbers, and visualizations into them depending on the cognitive styles of the target audience (Davenport *et al.* 2001).

To sum up, insights can be defined as the outputs of the predictive and prescriptive analytics models. The perceived value of the insights appears to depend on factors like disclosing model assumptions, using comprehensible models, and communicating the results in an easy-to-understand fashion. It should be noted that the difference between disclosing modeling assumptions and using comprehensible models is not clear, as the assumptions often depend on the employed model. In the empirical part of the thesis, we will attempt to gain a better view what affects the value of insights and what is their relevance in the process of analytics.

2.7 Insights are used to make and implement business decisions

Finally, according to our preliminary framework, the final stage of the analytics process consists of making and implementing the decision. As stated above, some authors like Liberatore and Luo view decision-making and implementation as one stage, whilst some, for example Davenport *et al.* consider it as two stages (Libetore and Luo 2010, Davenport *et al.* 2001).

The decision-making and implementation stage of analytics is an indispensable processual stage because insights have little value unless managerial actions can be taken based on them (Liberatore and Luo 2010). The view is shared by Vidgen *et al.* who have reported as a key finding of their study that generating value out of analytics and employing analytics to make business decisions are paramount analytics concerns for the experts and leaders of the field (Vidgen *et al.* 2017).

In contrast to the nature of insight, human decision-making processes have received ample attention in economics/business literature, and several Nobel Prices in Economic Sciences have been awarded to the researchers of the field. In literature, a distinction is often made between intuitive and analytic modes of thinking and decision-making (Kowalczyk and Buxmann 2015), which are both hereafter referred to as intuitive thinking and analytical thinking.

Intuitive thinking refers to fast automatic thinking that is based on one's own personal experience and loose heuristic rules. Intuitive thinking has been found to cause to systematic errors in decisions. (Tversky and Kahnemann 1974) According to Kowalczyk and Buxmann higher levels of intuitive thinking are also connected to low quality business decisions, while good business decisions are more often made as a result of analytical thinking. (Kowalczyk and Buxmann 2015)

Analytical thinking refers to systematic information collection and relying on rational analysis to make the decisions. Analytical thinking has been found to contribute to the good quality of business-decisions (Kowalczyk and Buxmann 2015). It should be borne in mind that even analytic thinking is always only boundedly rational because a human decision-makers cannot go through and take into account all case-pertinent information (Simon 1955).

Following Kowalczyk and Buxmann's line of thinking, we deem that AA solutions should increase the proportion of analytical thinking in decision-making and correspondingly reduce the amount of intuitive thinking (Kowalczyk and Buxmann 2015). The same approach has also been advocated by Power (Power 2016). As said, Kowalczyk and Buxmann have authored an article setting out several methods for increasing the role of analytical thinking in analytics. However, their methods relate to all of the stages of the analytics process, not only the (second to) last stage of decision-making, whereas now we are concerned with improving only the decision-making stage.

Davenport *et al.* have proposed improving the decision-making part of the process by organizing decision-audits, that is, by making the decision process transparent and documented, and assessing the quality of the decisions after the decision outcomes have been realized. (Davenport *et al.* 2001) This approach would appear to provide a *post hoc* feedback mechanism enabling the identification and weeding out of unwanted intuitive decision-making patterns.

Davenport *et al.* have duly pointed out that decision-making and its implementation can take place on different levels of the organization, managers making the decision and employees being in charge of the final implementation (Davenport *et al.* 2001). Also, Seddon *et al.* highlight that in absence of real-world implementation, the results of analytics initiatives are not realized (Seddon *et al.* 2016). Due to these arguments, we will modify our preliminary framework, and henceforward consider implementation of decisions as its own stage.

Davenport *et al.* also suggest that the quality of implementation may be assessed on the basis of behavioral, processual, and financial outcomes. The first one refers simply to whether the decisions are in fact implemented by the employees who do the work. Second, process changes imply changing the continuous way of working. Instead of implementing a one-off decision, the entire way of working is changed, by for example, by automating a previously human-driven process. Finally, the Davenport *et al.* propose measuring the financial outcomes of the analytics initiatives, and arguing cost savings being easier to measure than revenue increases. (Davenport *et al.* 2001)

To the knowledge of the writer, what has not been covered in analytics literature, is the accommodating time-value of quick analytics implementation into implementation's financial success metrics. For example, Shields has advocated focusing on quickly achievable wins in analytics initiatives but has not provided any theoretical groundings to back up his recommendations (Shields 2016). Further, LaValle *et al.* appear to suggest that the easiest problems having the same financial impact should be addressed first (LaValle *et al.* 2011). However, the financial advantages of quick implementation of a single analytics initiative appear not to have been discussed in literature.

The old adage absence of evidence does not equal to evidence of absence should not be forgotten. The reader is explicitly warned that the following line of reasoning may already have been found by other authors.

Advanced analytics solutions provide optimization-based recommendations on possible future courses of action. Optimization looks for an optimal solution to a specific problem. To my knowledge, the outcomes of the optimization do often ultimately reach some maximum or minimum value, with improvements requiring relatively more effort to achieve. If these premises hold, there should be value in realizing the benefits of analytics initiatives quickly by implementing "quick and dirty" working solutions instead of developing and launching the perfect solution after a long development. As a result, the net present value of the implementation would be higher. This hypothesis will be tested in the empirical part of the thesis.

2.8 Concluding remarks on the process view of analytics

In the beginning of the thesis we laid out the research questions. To recap, the two first research questions were: (i) are there any intrinsic characteristics of advanced analytics solutions that separate them from analytics; and (ii) which factors affect AA solutions' implementation success. This concluding sub-chapter sums up the answers that we have so far formulated to the two research questions and also sets out testable hypotheses for the empirical part of the thesis.

Our preliminary framework adopting a process view to analytics provided us with a robust instrument for exploring analytics in a structured manner and enabled us to connect various complementary findings from analytics literature in their proper contextual places. By employing the preliminary framework, we found that the intrinsic characters of analytics are that it can be regarded as a process with clearly identifiable stages. In short, these stages are of obtaining and preprocessing data, transforming it into insights, using the insights to make business decisions, and implementing the decisions.

By analyzing relevant literature, we understood that data can be viewed either as an enabling factor or a processual stage of analytics. Further, data is not an isomorphous topic but the concept may refer to structured or unstructured data, both requiring differing treatment methods. Appropriate limitations to the scope of the study were made, according to which this thesis only deals with analytical methods used in a structured data environment.

Second, the transformation of data into insights is not always a linear process. The cross-tensions between the SoftOr-oriented views regarding analytics as a tool for loosely defined problem solving, and on the other hand more traditional OR approaches viewing analytics in more rigid terms, were settled by acknowledging the importance of human judgment, experimentation, and iteration when transforming data into insights. Human judgment, experimentation and iteration are often needed before the best possible transformation methods are found. A reasonable assumption is that the skill in using human judgment, experimentation and iteration are a factor affecting AA implementation success. This assumption will be tested in the empirical part.

The preliminary framework also enabled us to answer the second part of the first research question: what separates advanced analytics from "plain vanilla" analytics. By focusing on the contents of the second processual stage, i.e. transformation of data into insights, we established that the methods used in the field of analytics have differing orientations. Whereas descriptive analytics takes a look at the past, predictive and prescriptive analytics are both focused on the future. Predictive and prescriptive methods attempt to predict future courses of events, and give recommendations on optimal courses of action. Further, the two are practically inseparably connected to each other because prescriptive analytics often utilizes the results of predictive analytics. Lastly, the same terminological choice of labeling predictive and prescriptive analytics as advanced analytics has also been made in earlier analytics literature. In the empirical part, we will find out whether this distinction makes sense to industry professionals.

Resultingly, we are able to define advanced analytics as follows: In this thesis analytics means the process of iteratively transforming data into insights through analysis for the purpose of making better business decisions. Advanced analytics refers to computer-supported decision-making methods that generate predictions on the future courses of events and provide recommendations on how to allocate available resources efficiently taking into account the objectives, and constraints of the operating environment.

Third, we found out that the perceived quality of the results of the transformational process, i.e. *insights*, may vary. Some insights are more valuable than others, and the experienced quality of insights appears to depend on several factors like making the background assumptions clear to decision-makers, using comprehensible models, and communicating the results to decision-makers in an easily digestible fashion. In the empirical part, we will try to obtain more information on how the quality of insights affects the process and how it can be improved.

The adverse influence of managerial intuitive thinking in the second-to-last stage of business decision-making was noted. Based on literature advanced analytics solutions should constrain managerial intuitive thinking by, for example, providing *post hoc* control mechanisms enabling after-the-fact-decision audits. In the empirical part, we will attempt to cast light on improving the decision-making process.

Lastly, we noted that decision-making and implementation are separate but interconnected stages. Managers' decisions are normally implemented in the lower rungs of the organization. Research suggests that the success of implementation can be evaluated on the basis of financial, processual, and behavioral outcomes of the decisions. However, we have not so far identified any concrete measures on improving the implementation success of analytics. In the empirical part, we will also attempt to identify what kind of means are available to improve the implementation stage, and also verify whether the hypothesis of time value of quick implementation withstands the test of reality.

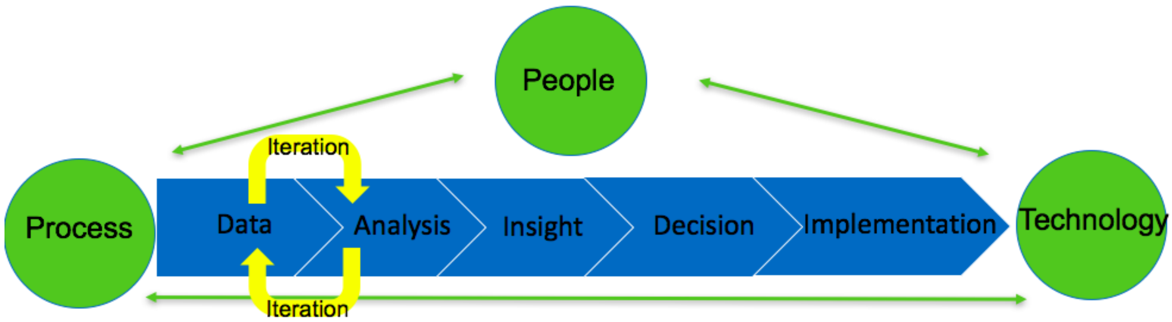

Figure 5: The revised process model of analytics and success factors

2.9 Analytics success factors

2.9.1 Purpose of this chapter

In this chapter, the reader will be introduced to the factors, which literature has found to affect how the business value of analytics is realized. This sub-chapter has two goals. First, it is useful have a broad view on the factors facilitating or hindering the adoption of analytics in organizations. Second, by defining the context of analytics in a detailed fashion, the critical assumptions behind this study can made explicit for the readers' assessments. As stated above, technological aspects of analytics are taken as given in this thesis, and are described to the reader solely for providing necessary background knowledge.

2.9.2 Technology

The first factor contributing to the success of analytics initiatives is (information) technology (Shanks and Bekmamedova 2012), encompassing both hardware and software. Mortenson *et al.* and Kohavi *et al.* have argued that technological progress in the field of data processing played a vital role in the growth of analytics (Mortenson *et al* 2010, Kohavi *et al.* 2002). Liberatore and Luo set forth that increasing computing power, better databases, and reducing storage costs have greatly increased the amount of data that can be collected, stored, and processed (Liberatore and Luo 2010). Seddon *et al.* found that enabling technologies, like hardware and software contribute to business value of analytics (Seddon *et al.* 2017). Thus, most analytics authors share the view, albeit with differing accentuations, that technology indeed is a critical success factor of analytics

Next, we will take a deeper dive into hardware and software on which advanced analytics capabilities are built, since the use of advanced analytics requires access to sophisticated information technology resources.

2.9.3 Hardware

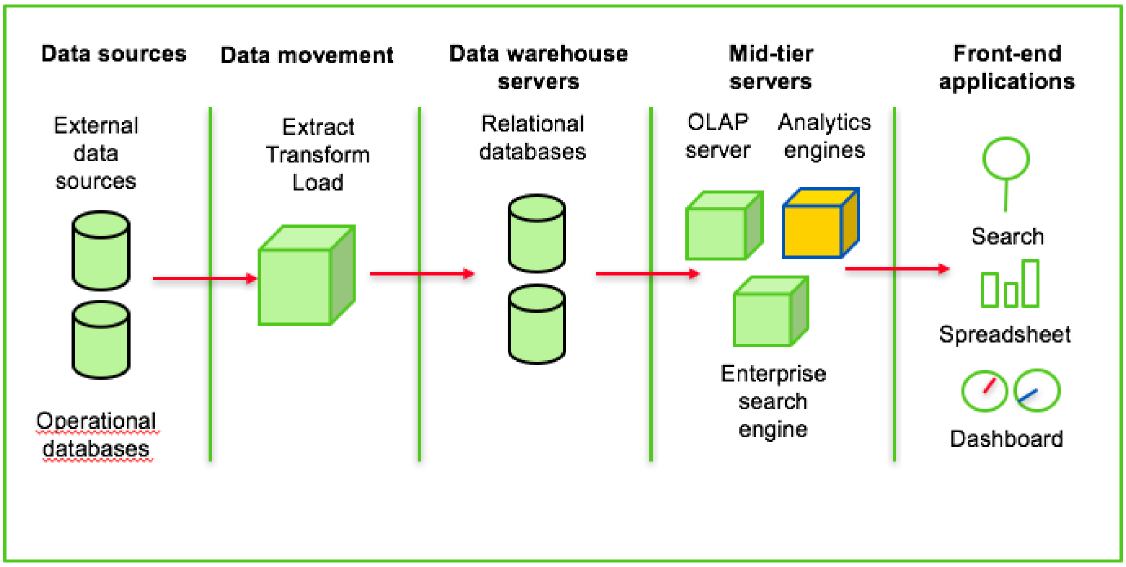

Figure 6: The technological infrastructure of analytics (modified from Chaudhuri *et al.* 2013)

The first layer of the required technology architecture – known as the technology stack – consists of the data sources providing the source material for the analysis. The data originates from different sources of varying quality, such as departments' operational databases and vendor databases. The data is not always stored in uniform formats and it may also contain missing and/or misleading entries. Thus, the data needs to be extracted from the original sources, transformed and only then loaded into the target database(s). The preprocessing steps are called ETL, which refers to extraction, transformation, and loading. (Chaudhuri *et al.* 2011) ETL forms the second layer of the technology architecture and can either be performed manually or in software. Often the ETL measures require considerable human effort (Liberatore and Luo 2010). After the data has been cleaned, it is typically aggregated in a relational database server called a data warehouse. Data warehouses have to be capable of completing SQL queries effectively and quickly. (Chaudhuri *et al.* 2011) Some authors like Gandomi and Haider refer to the data sourcing, ETL and subsequent data warehousing collectively as data management (Gandomi and Haider 2015).

The data warehouses are complemented by mid-tier servers in charge of performing specialized tasks. These specialized tasks include, for example, providing responses to online analytic processing (OLAP) enquiries. OLAP servers perform elementary analytics operations like filtering, aggregation, and pivoting. (Chaudhuri *et al.* 2011) The OLAP operations falling under the field of descriptive analytics provide less sophisticated insights than what advanced analytics does.

What is interesting for the purposes of this thesis are the data mining engines, on which the advanced analytics solutions (models) are run. The data mining engines can perform much deeper analysis than the OLAP servers. According to Chaudhuri *et al*. the using the data mining engines on a large scale is challenging. The analytics engines operate on data subsets drawn from the data warehouse. The analytics engines calculate the data's key statistical characteristics on the basis of which the predictive models are built. (Chaudhuri *et al.* 2011)

Chaudhuri *et al.* suggest that the challenging phase follows when the predictive models are integrated back into the operational databases. Special attention has to be paid to the data transfers that take place between the data warehouse and the data engine, and scalability of the solution. One approach is to integrate the analytics into the back-end data warehouse to overcome the data transfer and scalability issues. Traditionally the data mining technology has been packaged separately by companies offering statistical software. (Chaudhuri *et al.* 2011)

Lastly, what is visible to the end-users are the front-end applications. The front-end applications are interfaces used to send commands to the other layers of the stack for making searches, ad hoc queries, and showing the resulting information on spreadsheets and dashboards Chaudhuri *et al.* 2011).

The technology on which the advanced analytics solutions run on is seldom built from scratch. Most often, pre-existing legacy IT architecture defines what kind of analytics solutions can be implemented. Legacy systems and outdated IT architecture may prevent the combination of data from multiple sources or not permit the continuous processing of data, which makes the employment of analytics more difficult (Barton and Court 2012). In the same vein, Bose has argued that progress in technology has to be balanced with systems that are known to work (Bose 2009). In sum, integrating the advanced analytics components with existing systems cannot be overlooked because implementation of analytic insights in operations creates the business value (Kohavi 2002).

2.9.4 Software

According to Liberatore and Luo's suggestion, the availability of analytic software tools has driven the use of analytics in organizations. To back their claim, they argue that analyzing data is nowadays a lot easier because software suites containing graphical interfaces and many readily available statistical methods to process the data are included in the software by default. (Liberatore and Luo 2010) Chaudhuri *et al*. have similarly argued that the traditional way to employ analytics has been the use of prepackaged software (Chaudhuri *et al.* 2011). Whereas the above-mentioned authors highlight the role of off-the-shelf software's, like SAS, Cognos, and Fair Isaac's role (Liberatore and Luo 2010, Chaudhuri *et al.* 2011), Kohavi *et al.* claim that it has been the custom-made industry-specific applications explain the success of analytics. (Kohavi *et al.* 2002)

What seems to not have been sufficiently covered in the literature is the role of open source software in advanced analytics. Open source software programs and languages with versatile and quickly developing modeling add-on modules, such as R and Python, have also been widely employed by analytics professionals in various AA contexts.

Regardless of which interpretation one prefers – the ones suggested in literature or the one suggested by the author of this thesis or some combination thereof – it is beyond doubt that the better availability of analytics software has indeed sped up the use of analytics in business by providing the tools to create the AA solutions.

2.9.5 Process orientation to operations

Second, process orientation has been suggested to contribute to the success of analytics initiatives. In a process-oriented organization, the tasks that constitute a process are well defined and accurately measured, which facilitates their quantitative analysis, i.e. the use of analytics. (Liberatore and Luo 2010, Shanks and Bekhmadova 2012, Davenport *et al*. 2001) Similarly, using somewhat more general terminology, Seddon *et al.* argue that the extent of embedding evidence-based decision making into an organization's processes affects analytics success (Seddon *et al.* 2016).

In practice, information used to improve the process should be readily available, timely, and relevant (Shanks and Bekmamedova 2012). Process-orientation also requires tightly coupling and embedding analytics within the business processes (Shields 2016, Shanks and Bekmamedova 2012, LaValle *et al*. 2011). Based on first appearances, it seems undeniable that the extent to which the organization’s members view the business as a process with its own metrics and controllable variables affects the success of analytics initiatives.

Even though considering process-orientation as a success factor seems to enjoy broad support in literature, the view is subject to criticism on theoretical grounds. It seems that almost all of the elements of process orientation mentioned in literature can be attributed to either technology or people. First, the decision of defining the goals of the analytics process and how the process is measured is made by people. Second, the availability of information and embedding technology into the business processes are arguably attributable to both people and technology. Information can be either communicated by people of by technological means. Additionally, if analytics is embedded in a decision process, it must be either technologically incorporated in the decision-support solution or taken into account by human stakeholders of the process. Thus, process orientation appears to be a theoretical construct, which may not create any extra value.

Thus, for the purposes of theory development, it seems reasonable to eliminate process orientation as a conceptual category containing the above elements, and instead to move the success factors attributable to process under the two other main categories, technology and people.

2.9.6 People

Many authors have concluded that people affect the success on analytics in many ways and on all levels of organization. Literature contains a wide range of opinions on top and middle management’s, analytics professionals’, and rank and file employees’ roles in the successful adoption of analytics. Depending on the author, said stakeholder groups can have either a positive and/or a negative impact on analytics implementation success. Some of these points of view appear to be relevant in the context of this study.

Certain authors, like Liberatore and Luo have highlighted the role of managers. In their opinion, not backed by empirical evidence, the latest generation of managers, which was the first one to grow up with computers and recognizes the value of IT in business management, has finally reached executive positions and had a favorable impact on analytics initiatives. (Liberatore and Luo 2010)

Seddon *et al.* came to a similar outcome as a result of their analysis having the questionable empirical backing. In their view, analytics leadership, that is top management's commitment to advance and prioritize analytics initiatives as well as allocate resources to analytics projects, contributes to the success of analytics (Seddon *et al.* 2017). The fact that managers have been found to have an important role in analytics implementation is hardly surprising. We defined above that one stage of analytics is decision-making, and decision-making is the realm of managers. Therefore, following basic predicate logic, managers ought to be involved in analytics.

Further, many authors share the view that one of the greatest problem that analytics-driven decision-making culture faces is the managers' uncomfortableness to make decisions on the basis of the models they do not understand (Viaene and Van der Bunder 2011, Kowalczyk and Buxmann 2015). As seen above, in cases of lacking analytical understanding, managers tend to make decisions on the basis of their intuition, which has been claimed to lead to worse decisional outcomes in structured contexts (Kowalczyk and Buxmann 2015).

Some organizations have even attempted to tackle the problem of lacking managerial knowledge by providing analytics training to their managers. Ranbotsham *et al.* found in their survey of 2,719 business managers, executives and analytics professionals that approximately one half (49%) of the surveyed organizations have made an explicit effort to train their managers in analytics (Ranbothsham *et al*. 2015).

Regarding the analytics subject-matter experts, Viane and Van der Bunder have advocated they open up the metaphorical black box of analytics for narrowing the informational gap between the analytics experts and decision-makers (Viane and Van der Bunder 2011). According to Kowalczyk and Buxmann, analytics experts can mitigate decision-maker non-comprehension problems by limiting the number of models used, and by being transparent with the methods employed. The methodological transparency refers to using models decision-makers understand, providing them with information on the modeling assumptions, and using business subject-matter experts to define the goals of analytics projects. The authors also encourage the use of so-called analytics integrators, whom have good skills in explaining and visualizing the results of the analytics process. (Kowalczyk and Buxmann 2015) Also, easy interpretability of the outputs of the analytics process has been supported in literature (Davenport *et al.* 2001).

Regarding the end-users of analytics solutions, Seddon *et al*. have found that functional fit of analytics tools is important. Functional fit means that an analytics tools enable quick access to data and effective and efficient analysis of data. With respect to the end-users' psychological characteristics, Seddon *et al.* set out that overcoming organizational inertia is critical for analytics success. The authors define organizational inertia as the extent of motivation to learn, use and accept the new system. (Seddon *et al.* 2016) These findings are consistent with the ones made by Davenport *et al.* who argued that unless the employees implement the decisions made during the process, the benefits will not be realized (Davenport *et al*. 2001).

2.9.7 Conclusions on the factors affecting the success of analytics implementation

In the beginning of this chapter, we briefly described the technological underpinnings of analytics. The description was provided to acquaint the readers with the technologies underpinning analytics.

Second, the importance of adopting a processual approach to business operations was clarified. The better defined the processes, their inputs and outputs are, the better the chances for analytics initiatives success are. However, under closer scrutiny, we found that the elements of processual orientation can more naturally be attributed under the thematic categories of people and technology.

Lastly, people were found to contribute to the success of analytics initiatives on top-management, managerial and employee levels. In sum, the success of the analytics initiatives may be improved by making sure that analytics solutions are available, and easy to use and understand to all involved stakeholder groups.

Based on the contents of this chapter, we can conclude that the measures taken to improve the success of analytics initiatives can be connected to both the stages of the analytics process and the persons in charge of implementing a stage of the process. For example, in the processual stage of insight generation we found that some authors had recommended producing easily understandable insights, whereas other approached the matter from another angle, and allocated the responsibility for producing comprehensible insights this to analytics experts. Similar findings seem to concern all the stages of the analytics process.

Since, we notice that the literature starts to offer similar or even the very same answers to our second research question, that is, what influences the success of analytics initiatives, it is the time to move to the empirical part of this thesis to test our theoretical findings.

3 DATA AND METHODS

In the third chapter, the structure of the empirical part is explained to the reader. First, the context of the case study and the data sources are described. Second, the methodological background is set out. Lastly, the trustworthiness of this study is discussed.

3.1 Method selection

This study was conducted using the single case-study method and adopting elements from the systematic combining approach introduced by Dubois and Gadde (Dubois and Gadde 2002), all elements of the method falling under the umbrella of qualitative research. The details of the method are set out below. The main reason for conducting a qualitative study is that quantitatively oriented research on advanced analytics success factors would have required extensive access to a large number of target organizations for collecting the needed data. Such arrangements were unfortunately inaccessible to the author.

Case study method has been characterized as a good methodological choice where established theory is scarce. As stated above, established theories on advanced analytics implementation success factors are not numerous. Current theories appear either to deal with “plain vanilla” analytics or lack rigorous empirical support. Thus, it looks like there would still be room to validate the findings of earlier analytics literature in the context of advanced analytics and perhaps to even develop new theory.

According to Eisenhardt, case study is an approach for understanding the dynamics that are present within individual settings. A case study can be divided into stages. The case study process starts from developing *a priori* constructs, which guide the collection of data (Eisenhardt 1989). This means that the researcher should develop preliminary ideas on what kind of a phenomenon is investigated and which variables have causal effects on the phenomenon’s end-states. As regards the preliminary ideas, Eisenhardt claims that developing theories and relevant variables should be avoided in the beginning of the process to create space for unhindered theory creation (Eisenhardt 1989). In contrast, other authors, such as Dubois and Gadde advocate relying on theory to guide the data acquisition and analysis to avoid indiscriminate wide-scale collection of data (Dubois and Gadde 2002).

In this thesis, the more structured approach proponed by Dubois and Gadde is adopted in order to focus the research efforts. Also, taking into account that earlier theory on analytics exists, a primarily theory bound, deductive approach, in which the main emphasis is on confirming the findings of earlier theory seems appropriate. The development of preliminary theoretical constructs was performed during the literature review and the subsequent synthesis. If needed, the preliminary framework will be adjusted and updated in the light of the findings made during the research process.

Dubois and Gadde argue that during the research process the preliminary theory may be have to be adjusted on the fly if unanticipated but relevant findings emerge. Two mechanisms are suggested for adjusting the theory: matching, and direction and redirection. Matching refers to "*going back and forth between the framework, data sources, and analysis*" during the research. Direction and redirection entails steering the research into totally new directions, if the data requires so. (Dubois and Gadde 2002)

In this thesis, the opportunities for exercising matching, and direction and redirection were kept open by structuring the interviews in a manner that enables iteration and redirection of questions. The interviewees were first asked open questions, and only thereafter were the hypothesis postulated by the framework tested. The purpose of such "from general to specific approach" was to enable the researcher to ask complementary questions and to discover new theory before the interviews and the interviewees' preconceptions were "tainted" with superimposed theoretical concepts of the preliminary framework. Furthermore, where it seemed beneficial, questions probing emerging themes were often asked.

3.2 Case company selection and background

In case studies, the aim of selecting a target case entity is replicate earlier theory or to permit the development of emergent theory (Eisenhardt 1989). The aims of the case study method were the key driver in the selection of the case company. The case company selected for this thesis is Sellforte Solutions Ltd, a small Finnish limited liability company (Sellforte) established in May 2017.

The case company provides advanced analytics solutions for multiple retail customers and employs seven persons as of August 2018. The firm's advanced analytics solution can be used to predict and optimize promotional pricing and to direct marketing investments, both in order to improve a retailer's profitability. The company has been able to develop its analytical methods and technology infrastructure from scratch without any constraints imposed by legacy systems or predetermined methods. The unconstrained technological and institutional solution space has enabled the company to experiment with novel analytical methods and technological innovations. The company's technological and analytical freedom stands in a stark contrast with the control interviewee's organization, the practices of which are constrained by strict regulatory requirements.

As the case company operates in the relevant industry, the prospects of testing existing theory seem to be good. Further, the company is young and has taken an innovative approach to analytics. As the firm is be constrained by old technological and methodological choices, the prospects for developing new theory appear to be high. It should be noted that the study has not been commissioned by the case company. Neither does the author have an official role in the firm.

3.3 Data

This study uses semi-structured interviews as the data source of this thesis. Semi-structured interviews were selected as the data collection method because they enable collecting deep and rich context-sensitive information on the research questions. Further, semi-structured interviews do not limit the discussion to pre-determined topics but the interviewees are able to express their own interpretations on other themes relating to the research questions.

Before the interviews were conducted, an interview protocol was drawn to ensure reliability of the study. The protocol is appended as **Appendix 1** of this thesis. The appendix sets out the interview questions and is based on the hypothesis and themes developed during the earlier framework synthesis. In the interviews, the protocol was mainly used to ensure that all the relevant topics were covered. The interviews were held in Finnish, recorded, and their contents transcribed, all of the above with the interviewees' consents.

A total of four persons were interviewed. Three of the interviewees work in the case organization. The remaining control interviewee shares all the same educational and professional attributes than the case organization participants, but works in a differing organizational setting.

All of the interviewees purposely share similar educational and professional backgrounds. First, as regards the professional background, all interviewees are heavily involved in the development and implementation of advanced analytics solutions. Further, all interviewees share almost identical educational qualifications. Two of the case organization interviewees have received their master's degrees in engineering from Aalto University School of Science, Department of Technical Physics and Mathematics, having systems analysis as their major. The third interviewee has received a master's degree and an additional D.Sc. from the same department with a physics major.

The control individual is also an engineer by trade, having graduated from the same university than the four other ones, although with an industrial engineering major. However, the control individual has studied systems analysis as a minor, together with the demanding long mathematics curriculum. Therefore, when it comes to setting up advanced analytics solutions and developing the required mathematical models, the educational differences between the interviewees seem to be almost insignificant.

Also, with respect to business acumen, like results presentation and communication skills, the interviewees appear to have comparable skillsets. Two of the case organization interviewees have several years of analytics experience from the private sector and so has the control individual. However, one of the case firm's employees has a background in academia.

In contrast, the organizational settings differ considerably. Whereas the three case organization interviewees have had a great liberty in building their advanced analytics solution from the ground up, the one built by the control interviewee's team in the pensions insurance fund had to be integrated tightly into his employer's legacy computer systems and existing investment and risk management processes. Also, the control individual works in a strictly regulated industry, pension insurance, which limits the selection of models used in analytics solutions as well as the contents of the recommendations they give. The control interviewee's organization employs approximately 500 persons and has over 20 billion euros of assets under management.

Interviewee	**Education**	**Responsibilities**	**Experience**	**Interview details**
Case firm CEO Referred to as Interviewee 1	M.Sc.(eng.), systems analysis	Project management, result communication, big picture modeling	1 year case firm, 5 years management consulting	27 April 2018,
Case firm Chief Technology Officer Referred to as Interviewee 2	M.Sc.(eng.), systems analysis	Technical infrastructure, UI development	1 year case firm, 5 years industrial company	27 April 2018,
Case firm Chief Science Officer Referred to as Interviewee 3	D.Sc.(tech.) M.Sc.(eng.), engineering physics	Technical infrastructure, modeling	1 year case firm, 5 years academic career	27 April 2018,
Control interviewee – Head of Quantitative Investment Systems, Finnish pension insurance fund Referred to as	M.Sc.(eng.), industrial economics	All of the above	3.5 years financial market institutions, 1.5 years pension insurance fund	3 May 2018.

Interviewee 4				

Table 1: Table of interviewees

3.4 Analysis of data

The interviews, all lasting approximately 45 minutes, were first transcribed manually, each interview producing on average ten pages of text. In the second stage, each research hypothesis developed in the theoretical part of the thesis was written out. That stage was followed by a read-trough of each interview transcript. All the interviewee's statements relating to a particular research hypothesis were written under the relevant hypothesis. After this, it was counted how many interviewees had expressed a view relating to the thesis to assess the strength of available empirical evidence. Thereafter, it was analyzed whether the answers confirmed or contradicted the research hypothesis. After this stage the consistencies and inconsistencies between the interviewees' statements were analyzed and reported. Finally, the emerging themes not included in the interview questions were observed, and the interviewees' opinions were collected and analyzed following the same procedure set out above. The method of analysis can be characterized as inductive, as the view on the nature of reality was constructed on the basis of individual observations.

3.5 Trustworthiness of the study

According to Lee, the case study method has four potential pitfalls potentially affecting the trustworthiness of the study. The pitfalls relate to the method's ability to (i) make controlled observations, (ii) make controlled deductions, (iii) enable replicability and (iv) enable generalizability. (Lee 1989)

By the first one Lee refers to the difficulty of eliminating the impact of alternative explanatory factors. Whereas in laboratory sciences and statistical analysis, the influence of alternative explanatory factors can be adequately taken into account, similar controls are not available in case studies. (Lee 1989) The findings of research could thus – in theory – be explained by some unobserved factors. In this master's thesis, the hypothesis regarding the key explanatory factors are based on empirically tested findings of earlier analytics literature. Further, in the hypothesis development, attention was paid to identifying alternative explanatory factors. Lastly, the control interviewee working in an altogether different environment than the case company's employees serves as the last line of defense in identifying alternative explanatory variables. It cannot be entirely ruled out that the final findings of this thesis are explained by some unobserved factors but the risk of it seems to be mostly theoretical.

Second, Lee argues that the second pitfall of the case study method is that controlled deductions are hard to make (Lee 1989). Unfortunately, Lee gives very little guidance on what making controlled deductions means. In the author's view Lee means that the theory through which the facts are analyzed can be faulty and thus produce incorrect conclusions from the given facts. In the context of this thesis, making false deductions could imply that the presence of success factors in an analytics project would not contribute to the success of the project. This seems very unlikely, since in literature all of the success factors have been found to work.

Third, Lee points out that case studies are hard to replicate (Lee 1989). Non-replicability seems to be the key issue affecting the trustworthiness of this study. The facts of this study were collected from a single advanced analytics service provider offering its services to selected customers. The results are likely to be slightly different even if the same research was conducted again in the same environment due to the learning that has occurred during the subsequent projects. However, due to the general nature of the theory, the key findings would probably still apply.

Fourth, Lee has pointed out that the results may not be generalizable, i.e. apply in other similar contexts. Precautions were taken to improve generalizability by triangulating the findings by using the control interviewee. As stated above, the control interviewee works in a differing environment. If the findings from the interviews of the case organization members are similar those made in connection of interviewing the control interviewee, the results are likely to be generalizable.

4 FINDINGS AND DISCUSSION

4.1 Purpose of this chapter

The purpose of this chapter is to assess whether the research questions and the hypothesis constructed in the theoretical part of the thesis are supported by the views expressed in the interviews. This chapter lines out the research hypotheses in the same order than they were presented earlier. The interviewees' views on the research hypotheses are presented right after the description of each research hypothesis. If the interviewees expressed views that are relate a hypothesis in connection of answering to other research questions, these views are also presented.

The way of presenting and discussing the findings deviates slightly from the established practice of qualitative studies. Normally the findings are reported first and discussed in a subsequent chapter. The reason for combining the chapters was to eliminate unnecessary repetition of the hypotheses and empirical findings in the subsequent chapter. In the author's view, lining out individual quotes as representations of the interviewees' collective views is not representative and thus unnecessary. However, this approach requires that extraordinary measures are taken to avoid mixing the empirical results with the author's own interpretations.

The views presented during the interviews are separated from the author's views by adding the interviewee number after the opinions expressed by the interviewees. The interviewee numbers are set out in the table of interviewees. Furthermore, the text often attributes specific views to either interviewees or the author, by expressing clearly the source of information.

4.2 Is advanced analytics a relevant topic of study

The main premise of this thesis was that advanced analytics is separate subset of analytics. In the beginning of this thesis the term analytics was argued to refer to the transformation of data into insights, on the basis of which business decisions can be made and implemented (Liberatore and Luo 2010, Holsapple *et al.* 2014). It was further argued that advanced analytics is a subset of analytics. Advanced analytics was claimed to be limited to computer-based methods of analytics that predict and give optimization-based recommendations on suitable future courses of action.

We noted that the definition of “plain” analytics was well-established in literature (Holsapple *et al*. 2014), while multiple definitions had been given to the term of advanced analytics (Bose 2009), the most widely-used definition being similar to the definition adopted in this thesis (Gartner 2018, Barton and Court 2012).

According to the empirical findings of this thesis, the distinction between analytics and advanced analytics is not at all as clear as I initially claimed. On a general level, interviewees tended to recognize noteworthy differences in the specific methods used in the field of analytics. For example, two interviewees acknowledged that some methods used in analytics demand considerably more technical expertize than other methods (2, 3). To give an example, one interviewee regarded the use of predictive and prescriptive models being much more technically challenging than the use of traditional business intelligence reporting mainly consisting of data visualization tools (3).

When expressly asked to give definitions to terms analytics and advanced analytics, two interviewees labelled machine learning solutions as advanced analytics (2, 3). However, of the two, the other one admitted that the term has multiple meanings (2). The third interviewee’s understanding was aligned with the definitions adopted in this thesis. He suggested that predictive and prescriptive methods could be labelled as advanced analytics. Thereafter, the third interviewee stated that he cannot not come up with solid criteria for making a distinction between analytics and advanced analytics and thereafter questioned whether dividing analytics into plain analytics and advanced analytics makes any sense. The reason was that simpler methods can in some instances be more effective, than complex ones. Finally, he proposed that branches of analytics could be divided into descriptive analytics and prescriptive analytics (4). The second interviewee stated similarly that the key point in using analytics is which models work, not whether the model is something generally regarded as advanced (2). He also mentioned that the concept of advanced analytics can refer to many things, but pointed out that computers are needed because the difficult questions cannot be solved by resorting to thinking, pen, and paper (2).

On the contrary, the definitions given by the interviewees to plain-vanilla analytics were more uniform and resembled closely the ones used in literature (see e.g. Holsapple et al. 2014). When asked to define analytics, three interviewees gave almost word-for-word identical definitions with the one formulated in theoretical part of this thesis (1,3,4).

The first major conclusion of this study – unfortunate as regards the topic of the thesis – is that the definition of advanced analytics is very disputed amongst analytics practioners. First, the boundaries between descriptive and predictive methods are blurry. According to one interviewee, linear regression can be classified as a descriptive and predictive method (4). Second, the relationship between the various machine learning methods and predictive and prescriptive analytics is also unclear.

Lastly, in the author's view labeling some branches of analytics as advanced is arguably time sensitive. As the technology and available methods mature, the boundaries of analytics and advanced analytics may change. A couple of years ago, machine learning and AI were mostly of interest to statisticians and computer scientists but nowadays they are regarded as the most advanced branch of analytics amongst the general population. Consequently, instead of using the term advanced analytics, more established terms like descriptive, predictive, and prescriptive analytics should be used in scientific contexts.

4.3 The stages of analytics

4.3.1 General remarks on the processual view of analytics

According to the theoretical framework the stages of analytics projects are the pre-processing of data, data analysis, obtaining and communicating the insights, and finally making and implementing the business decision (Liberatore and Luo 2010). All of the interviewees appeared to think that the processual model developed in the theoretical part of the thesis reflects the characteristics of one-off analytics projects rather accurately.

However, the picture of the stages of an analytics project becomes more multi-faceted if certain new variables, quite often left unaddressed in literature, and thus not included in the preliminary framework, are introduced into the equation. The interviewees mentioned that a company's strategic considerations and outsourcing/internal production decisions may add new stages to an analytics project and/or affect the relative importance of the stages.

These findings are discussed next before moving on to the main topic – the detailed analysis of main stages of analytics.

4.3.2 The role of goal-setting in analytics

Three interviewees highlighted the importance of aligning the analytics initiatives with the company's strategy (1,2,3). Two interviewees mentioned that the effects of analytics project must be understood holistically and prospective trade-offs between the goals must be taken into account (1,2). As a practical example, one interviewee stated that minimizing marketing spending might initially reduce costs but later decrease the number of customers (1). Similar views have been expressed in literature. For example, Davenport et al. argued in 2001 that analytics projects should be aligned with a company's strategic goals (Davenport 2001).

Relating closely to the theme of strategic alignment, all of the interviewees stated that the business goals of the specific project must be well-defined before the work is started and continuously during the project. One interviewee claimed in this respect that the biggest mistakes and right choices in analytics projects are made already when deciding the goals of the analytics projects (1). Similarly, another one argued that the key precondition for a project's success is appropriate goal selection (4).

Some reasons were implicitly mentioned for why clear goals are needed. Two interviewees mentioned that modelling experts have a tendency to analyze modelling-wise interesting matters having limited business value if the business-relevance of the analysis is not ensured (1,3). An interviewee mentioned off the record that a leading Finnish retail company had allegedly used weeks of its data science team's capacity to optimize the food recipe recommendations give on its web-page. On the other hand, one of the keys to successful goal selection was argued to be having enough knowledge of statistics and analytics in order to be able to find suitable models with which to tackle the problems (4).

To conclude, the goal selection stage appears to be an important but fickle business. On one hand, modelling expertize is needed to recognize the problems, but too focused expertize may result in losing the bigger picture. A balance should be found between the two extremes.

4.3.3 Differing skillsets in internal production and external purchasing of analytics

Three of the interviewees expressed that purchasing (parts of) the analytics solution from a third-party service provider requires a different skillset than developing the solution in-house (1,3,4). The needed skills differ both at the purchaser and the provider sides. This seems to be a finding that has not been dealt with in earlier research.

One interviewee said that when the analytics solution is purchased from a third party providing the same solution for multiple customers, the external supplier must simultaneously with the execution of a single client project ensure that the analytics for solution is fit for multiple clients' use. Thus, the multiple-client setting appears to set additional requirements for the development of the solution. It was specifically mentioned that the models used in solutions offered for multiple customers must be more stable to enable the analysis of more variable data (3). Another interviewee claimed that when offering the solution to multiple customers, the underlying modelling techniques and the solution would still have to be of top-of-the-line quality (1).

The small interviewee pool is likely to prevent us from making any definitive conclusions on the potential specificity-generalizability trade-off. In other words, it remains unclear if the generalizability requirement for the methods used in the solution has an adverse effect on the preciseness of recommendations given by such solution. Future research should find out, if the scalability of the analytics solution to multiple customers' use is achieved at the expense of the quality of case-specific results.

4.3.4 Data acquisition and preprocessing

According to our preliminary framework data acquisition and successful pre-processing of data are prerequisites of success for an analytics project. Literature has found that availability of good quality data is important (Seddon *et al*. 2014, Kowalczyk and

Buxmann 2015). Confirming, yet more nuanced, findings were made during the interviews.

All interviewees explicitly mentioned that good quality data needs to be available. However, the availability of good quality data is not a guarantee of success for this stage. One of the interviewee pointed out that if the data depicts a very efficient market, such as the currency market, meaningful insights may not be found from it (4).

Two interviewees pointed out that an interesting connection exists between the data acquisition stage and the subsequent data analysis stage. They argued that the properties of the data also define what kind of methods can be used in the analysis (3,4). One of the interviewees further pointed out that the contents of the data acquisition and pre-processing stage may differ drastically between industries. According to him in some industries like retail plentiful internal data may be available. If internal data is available, the focus is on acquiring the data from internal databases, and cleaning and verifying it. However, if data is purchased from third parties, whom take care of the data acquisition, verification and cleaning, the required skills relate more to being able to assess the quality of the data rather than performing the preparatory measures. Third party data may be in some instances be very expensive so the purchaser needs to be aware of what he is purchasing. Further, the most exclusive data providers may only provide access to the best data, like credit card transaction records, to their most valuable customers such as successful institutional investors or hedge funds. (4)

These contextual differences of data acquisition have not been addressed in previous literature, which has often dealt with the data acquisition stage as a purely technical exercise without paying due attention to the challenges of using third party data. Especially, when more and more external data is becoming available all the time, the elements of successful data acquisition skills should be investigated in more depth in future research.

4.3.5 Transformation of data in insights

Our framework further suggested that the second stage of analytics consists of transforming the data into insights. In literature, it had been found that good human judgment as well as an experimenting and iterative approach are needed to develop the best possible method to analyze the data (Viane and Van der Bunder 2011).

The importance of having good experimentation and quick model iteration capabilities were found to be critical by all interviewees. Three other interviewees pointed out that it may not always be clear from the beginning where and how the business of the case value can be found (1,3,4). For that reason, the goals and applicable methodologies may have to be changed during the early stages of the project, stated two interviewees (2,3). According to one of the interviewees, plotting the data helps in identifying suitable model models for analyzing the data. For example, if linear regression is used, plotting the data contributes to spotting outliers and in visualizing the shapes of the distributions, which may both affect the usability of the model adversely (4).

Three interviewees also considered obtaining the preliminary results quickly important (1,2,3). One suggested that progressing fast in the beginning of the project contributes to the timely internal validation of modeling assumptions, which helps in avoiding extra work (3). Furthermore, realizing preliminary results quickly also enables the communication of results to the customer, ensuring that the end-customer and the solution builder share a similar view of the goals of the project. These communicational aspects were seen as important by two interviewees (2,3).

4.3.6 Insight generation

Third, our preliminary framework postulated that the transformational process produces *insights* as its output, which are used for an instrumental purpose of decision-making and/or action-taking in a business context in the next processual stage.

In literature, insights were defined as cause and effect relationships between things. Theory suggests that insights can be compared with each other, and the value of some insights was deemed to be higher than the value of other insights. (Seddon *et al.* 2016.) During the interviews, I found that certain methods of presenting the insights appear to increase the perceived value of insights, just as suggested in literature, and make it more likely that decisions are made on the basis of the insights. Intriguingly, some of the methods suggested for improving the value of insights were called into question by the interviewees.

In literature disclosing the modelling assumptions has been recommended to increase the value of insights (Viaene and Van der Bunder 2011). Similarly, literature advocates the use of comprehensible models (Kowalczyk and Buxmann 2015). The differences between disclosing the models' assumptions and using comprehensible models are not clear to the author. To the author's understanding, the specific models require specific kinds of assumptions to be made. Due to the apparent inseparability of the model and its assumptions, the benefits of disclosing modelling assumptions and using comprehensible models, should probably be viewed as one and the same issue.

One interviewee described an interesting contradiction relating to disclosing the modelling properties and assumptions that has not been mentioned in literature. He stated that if the analytics solution is provided by a third party, the solution's models' exact properties cannot be described to the clients in full detail because they can be reverse engineered (3). Such conflicting incentives of the environment may affect negatively the level of confidence the end-users have towards the solution. This issue should perhaps be researched further.

A second interviewee expressed a view that was very inconsistent with literature as regards the benefits of disclosing modelling assumptions. His remarks also apply to the alleged benefits of using comprehensible models. According to him, many end-users may not necessarily understand even basic mathematical and statistical concepts, such as p-values and confidence intervals, so trust in the results must be built more practical means, like appearing credible in the eye of the audience (1).

The evidence from a third interview supported the view questioning the audience's knowledge level. According to the interviewee, extrapolations based on historical results are often used as predictions of the future, even where such predictions are not statistically significant. The same interviewee also claimed that executives' lacking understanding of modelling may lead to outright harmful results. According to him, machine learning or “artificial intelligence” are currently regarded as silver-bullet-like solutions to all sorts of problems, which leads to misdirected investments (4).

Thus, the value of explicitly describing the modelling assumptions and details of the used models appears to depend to a great extent on the target audience’s technical sophistication.

Literature also recommends employing project management experts to communicate the insights to the target audience (Kowalczyk and Buxmann 2015). This hypothesis gained strong empirical support. Three interviewees, including the control interviewee, stressed that the analytics project manager creates benefits by being able to simplify matters and speaking in understood by the audience, which both contribute to bridging the gap in understandings of the analytics team and the decision-makers or users (1,2,4).

Lastly, literature suggests that packaging the outputs of the analysis appropriately, for example, by including an appropriate amount of text, numbers, and visualizations builds perceived trust in the results and increases the likelihood of their use in decision-making and implementation (Davenport *et al.* 2001). The interviews provided some support also for this hypothesis. According to one interviewee, the results must be presented in a form the audience understands which reduces the number of misunderstandings. If the analysis is sensible and logical, it can be presented in a comprehensible manner, argued another interviewee. According to him, if the analysis cannot be packaged in a comprehensible format, it may indicate that the analysis may not have firm grounds (4). Thus, the packaging of the insights matters in how the results are perceived by their users and decision-makers.

One thing not found in literature, but anyhow considered important by one interviewee related to the correctness of results. He highlighted on multiple occasions that the results of the analysis need to be correct. He argued that the results cannot contain any guesses or uncertainties. In case of such, the uncertainties must be communicated clearly to the end-users. It was claimed that unclarities and wrong results are especially harmful and degrade trust in the results of the analysis. (4) This finding could be also be researched further.

4.3.7 Decisions are made based on the insights

In the theoretical framework, the fourth stage of analytics is decision making. Literature sets out that decisions made on the basis of insights are essential in realizing business value. Some authors also claimed that analytics initiatives should aim to increasing the amount of analytic decision-making and correspondingly to reducing the role of intuitive decision-making (Kowalczyk and Buxmann 2015).

The interviewees did not explicitly use terms like reducing the proportion of intuitive thinking, but despite of that, the key message was the same. For example, one interviewee said that in when it comes to problems having an analytical solution, management should allocate more decision responsibility to the analytics experts, who have more subject-matter expertize and understand better the nature of the problems being solved (4). In essence, this point of view supports increasing the amount of analytics-based, i.e. is non-intuitive, thinking by re-allocating decision responsibility from the more intuitive-thinking executives to the analytics experts who are more analytically inclined.

Similar views were raised by another interviewee. He stated that decisions preceding the use of computerized analytics solutions are often based on gut-feeling, and are not based on a scientific approach (3). A more structured approach to problem-solving was claimed to add transparency to the decision-making. Thus, theory and practice appear to be aligned.

As a means for increasing the proportion of analytical thinking, literature (Davenport et al. 2000) suggests that decision processes should be transparent, documented and evaluated post-hoc. One interviewee found that analytics solutions improve decision quality by increasing transparency by the means of listing the possible decision options and related confidence intervals explicitly. According to him, good analytics solutions also track the made decisions, which enables their post-hoc assessment as well as the updating of the models and model parameters to work better in the future (3). Thus, the interviews provide clear but limited backing to the theory.

4.3.8 Decision implementation

According to our framework, the last stage of analytics consists of implementing the decision. According to Davenport et al., the success of implementation can be assessed on basis of three variables. First of all, the behavioural aspects of implementation refer to whether the decision is in fact put into practice. Second, in the optimal case, the decisions also lead to a constant improvement of the way of working. Lastly, the success of the analytics initiatives' implementation can be assessed on the basis of their financial outcomes. (Davenport et al. 2000)

The interviews provided some support for the view that the three variables are relevant metrics for assessing implementation success. In addition, the interviews cast light on how to influence these variables in a positive manner.

As regards the behavioral changes, the observed interviewee support was the very limited. Only one interviewee brought out indirectly the importance of putting the decisions into practice (1). This was surprising, since the interview questions asked what the relevant factors in assessing the success of the implementation stage are. The writer's interpretation is that the questions of actually implementing the decision may have been seen as too self-evident to be mentioned separately.

According to the interviewees, changes in the ways of working were also seen as a beneficial consequence of completing an analytics initiative by three interviewees (1,2). The benefits of using a computerized decision support solution were said to become to fruition partly when they were included as a part of the daily operating routines (1). In the writer's view, software products may be better than the deliverables of traditional consulting projects in creating lasting change in the ways of working. One interviewee claimed that if the results are presented on power point presentations instead of a software solution, the results must be simpler and they are also likely to disappear in the course of time (3). This is likely to reduce their value in the long run.

Third, all interviewees made it clear, unsurprisingly, that the success of an analytics initiative is often assessed on the basis of its financial outcomes. The interviewee working in the financial industry acknowledged that in the field of asset management the decisions are assessed to a large extent on the basis of their financial outcomes (4). Interestingly, assessing implementation success in asset management on the basis of realized financial results may not be as straightforward as suggested in literature. The interviewee stated that due to volatility of possible outcomes, an *ex ante* rational decision can in hindsight look like a bad decision because the outcome was negative (4).

On the other hand, where the amount of decisions made is large, the unexpected events are likely to represent an insignificant proportion of all decisions made, stated two interviewees (1,4). This opinion is consistent with statistical theory, according to which the sum of variances of independent random variables is proportional to the square root of the sample size. The evaluation of implementation success on the basis of financial outcomes was also seen as relevant in the field of retail.

In sum, where the decision is a one-off decision, such as a large capital expenditure decision, the quality of the initial decision should perhaps not be assessed solely on the basis of the realized financial outcome. In case the analytics solution is used continuously make small operating decisions, like retail pricing or asset management decisions, the average financial outcomes are likely to be a good criterion for judging decision quality. To the author's knowledge, this distinction has not been made in earlier literature.

Lastly, the author proposed as his own hypothesis that quick implementation of the analytics initiative would have time value. Time value referred to realizing the biggest benefits of the project quickly by using so-called quick-and dirty solutions, and iterating the solution afterwards to reap the smaller incremental rewards. Even though the interviewees emphasized the importance of quick implementation on multiple occasions, the financial time value of the quick implementation was not mentioned as a reason for the approach.

4.4 Success factors

4.4.1 Technological and people-related success factors

The second part of the theoretical review consisted of defining explicit success factors, the occurrence of which improves the likelihood of success in analytics initiatives. The success factors were claimed to be divisible in two groups, technological success factors and factors being attributable to humans. It should be restated that many such success factors have already been identified in connection of describing the stages of analytics.

In the theoretical part, the analytics solutions' compatibility with the company's existing technological infrastructure was regarded to contribute to the success of analytics projects. The technological infrastructure was argued to define to a great extent what kind of analytics solutions can be used. (Barton and Court 2012, Bose 2009).

The interviews did not provide support for this proposition. Three of the interviewees worked for a company offering a cloud-based analytics solution for their customers. The solution uses customer data but performs all the subsequent stages of the process in a proprietary cloud solution. It may be that due to the decoupling of analytics infrastructure from the customer company's legacy systems, the interviewees did not regard this matter as important. The technological constraints were not mentioned during the interview Neither was the matter raised by the control interviewee.

However, two interviewees expressed similar concerns on the effects of the available data on the subsequent analysis. They argued that the data defines to a great extent what kind of analysis can be executed. (3,4).

In the theoretical part of the thesis, people were defined as the second success factor of advanced analytics projects. Literature has found that four employee groups are particularly relevant for the success of analytics project. To recap, the four groups are the management, analytics project managers, analytics experts, and the end-user of the solutions.

In literature, top management was argued to contribute to the success of analytics initiatives by allocating sufficient resources to the project and by prioritizing the initiatives (Seddon *et al*. 2017). These premises were confirmed by three interviewees (1,2,4). As already stated above, one interviewee argued that the biggest mistakes are made and victories achieved when selecting the targets of the analytics initiatives, the management having a big responsibility in this respect (1). Similarly, the control interviewee claimed that the analytics experts play the biggest role in the success of analytics projects provided that management has manage to allocate the resources on the correct matters. Moreover, he expressed that the management must also define clear measurable targets for the initiatives (4).

One of the hypothesis developed in the theoretical part of the thesis was that the success of analytics projects in companies is hindered by the top managements limited understanding of analytics and modelling techniques (e.g. Kowalczyk and Buxmann 2015). As seen above in connection of analysing the decision stage, the interviews supported firmly the validity of the hypothesis.

Second, analytics experts play a key role in the success of analytics projects. Literature appears to divide the analytics experts in two groups. Project manager-like experts, whom supervise the project and work as bridge communicational builders between stakeholder groups form the first group. The second group of analytics experts consists of technical experts, whom prepare the data and analyse it (e.g. Kowalczyk and Buxmann 2015).

Literature suggests that analytics experts should bridge the informational gap between the experts and relevant stakeholder groups, like end-users and executives. According to literature, this can be realized by, for example, using comprehensible models and ensuring the transparency of analysis. Also, explaining and visualizing the results was recommended. (e.g. Kowalczyk and Buxmann 2015)

As argued above in connection of describing the decision-making stage, the benefits of using comprehensible models and disclosing the modelling assumptions depend on the sophistication of the target-audience. If the modelling understanding of the audience is limited, the value of the insights should be communicated by other means. In contrast, the importance of explaining and visualizing the results was acknowledged regardless of the characteristics of the audience.

In hindsight, after the interviews it looks like the ideal properties of the analytics subject-matter experts were not covered in a detailed enough manner in the theoretical part of the thesis, even though literature has dealt with the issue (Davenport 2006) On the basis of the interviews, it looks very likely that the analytics experts' skill-level has a great influence on the success of the project.

One interviewee maintained that the productivity levels of analytics experts can differ by an order of magnitude and be tenfold. Some analytics experts were described to self-determined and efficient, whereas bad analysts were claimed to focus on matters that are insignificant for the project. Recruiting the best talent was deemed crucial. (4)

The analytics subject matter experts were deemed to need excellent quatitative skills in addition to industry knowledge. Software development skills, such as version management, commenting code, and systematic testing, were considered important by two interviewees (3,4). Systematizing workstreams, and re-using existing knowledge were argued to create efficiencies in work by a third interviewee (2).

Lastly, literature has found that for end-users, the motivation to learn to use the analytics solution contributes to project success. This hypothesis was only indirectly supported by interview evidence. One interviewee's opinion was that the benefits of using the analytics solution are only realized when it is taken into daily use and the end-user start their continuous learning processes (1). It may well be that non-expert end-users should have been interviewed to gain more better insights on this hypothesis.

5 CONCLUSIONS

5.1 Research summary

5.1.1 The elements of advanced analytics

The first goal of the thesis was to find out what the differences between analytics and advanced analytics are. In literature varying definitions had been adopted for advanced analytics. The authors initially suggested that advanced analytics refers to computer-based methods of analytics that are used to predict possible outcomes and find optimal solution for the future scenarios.

However, on the basis of the interviews, it became evident that the distinction between analytics and advanced analytics may not be meaningful. The term advanced implies that advanced analytics is somehow more developed than normal analytics. The interviews made it clear that simple modelling methods can be more effective in solving particular problems than the methods we defined as advanced analytics. Thus, the terminology used initially may actually be misleading. Also, it is unclear whether machine learning-based models should regarded as analytics or advanced analytics. Lastly, it is unclear whether for certain statistical methods, for example linear regression, which appear to be suitable to both describe and predict the outcome of a phenomenon, should be regarded as a method belonging to analytics or advanced analytics.

Due to the term's misleading nature and lack of preciseness, the author recommends not using the definition of advanced analytics and instead using more established terms, such as descriptive, predictive and prescriptive analytics in scientific contexts.

5.1.2 The stages and success factors of analytics

The second research questions was to find out what the stages of analytics are, and which factors contribute to analytics initiatives' success.

The interviews provided ample support for the view expressed in literature, according to which analytics can be viewed as a process having distinct stages. The stages are acquisition and preprocessing of data, transformation of data into insights, making decisions on the basis of insights, and finally implementing the decisions.

In addition, the empirical findings support introducing goal selection as a new stage preceding the established processual stages. During the goal selection stage, the management chooses the areas of operational improvement, where business value can be created with analytical methods in alignment with the business strategy of the company. Analytics projects managers are needed in this stage to communicate what kind of goals can be achieved with the available technical and human resources.

In the next stage, good quality data needs to be obtained and preprocessed for the purposes of analysis. Literature had previously approached this stage as a technical data preparation exercise. The findings of this thesis suggest that the skills needed when using internal data differ from the scenario where data is acquired from external data providers. If the data is purchased, data procurement skills are needed to assess the quality, value, and price of the data. This is one of the main finding of the thesis. The contents of data acquisition skills should be investigated further in future research.

After data acquisition and pre-processing stage, the data is transformed into insights. According to the results obtained, experimentation and iteration skills are needed when selecting appropriate models. Data-analysts bear the main responsibility for executing this processual stage. The data analysts need to have a good knowledge of the relevant market and excellent mathematical and statistical skills. In addition, preliminary results must be obtained quickly to ascertain the suitability of the chosen methods for the selected purpose, and to confirm the quality of the results.

After the data has been analyzed, the insights created as a result of the analysis must be communicated to the decision-makers. Analytics project managers have the key role in this processual stage. Good analytics project managers can assess the knowledge level of the decision-makers, and present the results in a manner understood by the executives. Packaging the results appropriately, e.g by illustrating the results graphically increases the likelihood that they will be used in decision-making. However, one of the main findings of earlier literature was called into question. Research suggests that disclosing modelling assumptions and using easily comprehensible models increase the perceived value of the insights. The benefits of disclosing model assumptions and using comprehensible models appear to be dependent on the target audience's sophistication. In case that the audience does not have sufficient technical knowledge, providing too detailed technical information is unlikely to create any benefits.

The next stage of analytics is making the decision. The results obtained were similar to those of existing literature. Earlier literature suggests that in this stage, the goal is to reduce the share of intuitive decision-making and to increase the proportion of analytical decision-making (Kowalczyk and Buxmann 2015). This can be achieved by re-allocating decision responsibility to subject-matter experts on a lower level of organizational hierarchy. Further, the results suggest that computerized solutions increase the transparency of the process by explicitly lining out the decision options, their expected outcomes and related probabilities. This is likely to make the process more objective and efficient.

Lastly, the decisions must be implemented. The results confirmed that decision implementation is an important stage of the process, and creates the business value. However, any definitive results were not obtained on how to improve the results of this processual stage.

5.2 Practical implications

The major scientific contribution of this study was to collect and combine the main insights of earlier research under one set of hypotheses, which was also tested rigorously. Earlier efforts to systematize analytics had either employed a piece-meal approach to the topic (Kowalczyk and Buxmann 2015), been too holistic and thus overlooking the necessary details (Holsapple *et al*. 2014), or alternatively lacking firm empirical support (Liberatore and Luo 2010, Seddon *et al*. 2017). This thesis addressed all of the above problems.

The empirically tested processual framework of analytics has many practical benefits. The framework provides executives and managers with a coherent and detailed framework that helps in structuring their thinking process. The framework ensures that critical success factors are taken into consideration in all of the stages of analytics. In addition, the framework confirmed the existence of certain pitfalls, like the overreliance on intuition in executive decision-making, reported earlier by Kowalczyk and Buxmann, and provides means for avoiding these pitfalls (Kowalczyk and Buxmann 2015).

In contrast to the pre-existing models, the new framework highlights strongly the importance of the careful goal selection before commencing the analytics initiatives. The goals should preferably be defined in co-operation between the executives and project managers. This ensures that the analytics projects are executed in areas, where they create the biggest business value and that there are technical means available to achieve the goals.

5.3 Limitations of this study

The case study method has four potential pitfalls potentially setting limitations for the trustworthiness of the study. The pitfalls relate to the method's ability to (i) make controlled observations, (ii) make controlled deductions, (iii) enable replicability and (iv) enable generalizability. (Lee 1989) In the context of this study, making controlled deductions and the generalizability seem to be the most relevant pitfalls for forming scientific knowledge. By the former it was meant that the underlying theory, through which reality is observed, is faulty. Thus, wrong conclusions are made on the basis of the facts of the case. The latter referred to the fact that the results may apply in particular circumstances but not reflect the reality in general. (Lee 1989)

This study has two main limitations. The biggest limitation is that the interviewees all have a solid theoretical background in modelling but appear to have only limited exposure to the everyday use of the analytics solution. In this thesis, it was implicitly assumed that the interviewees view of the reality is correct. However, especially the end-users may think of the research questions in less mathematical and objective terms, and also have some important insights that are only observable to the end-users. Further studies should investigate the validity of the framework from the perspective of non-sophisticated end-users.

Second, all of the results of this study that emerged spontaneously during the interviews are not likely to be widely generalizable due to the size of the interviewee pool. The interviewee pool was relatively small, and consisted of four persons. In the authors view, the interviewee pool size was sufficient for verifying the hypotheses that were developed on the basis of literature. These hypotheses could be tested in all interviews. However, the emergent findings should be treated with care and require further validation. Due to their unexpected emergence, they could not be validated in multiple interviews but mostly reflected the individual views. For example, the hypothesis that acquiring data from third parties requires different skills than preparing internal data should be validated in a separate study. Due to the limited support of emergent findings, they should be treated as research hypotheses for future research but not as generally applicable results.

5.4 Suggestions for further research

During the research process two key areas for future research were identified. The first interesting research area is in the intersection of cognitive science and information system / analytics studies. Cognitive science researches how human cognition, i.e. understanding, works.

During the research, it became apparent that one of the bottlenecks in effectively employing analytics solutions relates to the gaps between the understandings of analytics experts and on the other hand executives and end users. The executives and end users may be able to fully benefit from the insights and decision alternatives produced in the analytics solutions because they do not understand the mathematical and statistical operations performed during the process. Thus, the decision alternatives cannot be effectively compared with each other because e.g. the expected outcomes and their confidence intervals are not fully incorporated in the analysis. This problem could perhaps be mitigated with the tools of cognitive science. Insights from cognitive science could potentially offer us with more detailed tools to analyze which features affect acting on the basis of non-understood insights that must be taken for granted.

The second set of interesting research question relates to applying process decoupling to analytics initiatives. In process decoupling, clearly separable stages of a process are identified. Thereafter, the stages of the process are assigned to parties that can perform the stages most effectively. (Metters *et al*. 2012)

Analytics initiatives have clearly separable stages, with different kinds of expertize required in each stage. For example, purchasing data from external parties requires a different skillset than analysing the data. Furthermore, presenting the results of the analysis requires other kinds of skills, like aptitude to visualization and modelling. Thus, it looks like the stages of the analytics process could be decoupled from each other, and assigned to best available internal or external providers. As we saw, such development is already taking place in a limited scope. Data collection and pre-processing can already be externalized. Moreover, sometimes the entire process of modelling and creating the insights has been externalized, as the example from Stockmann shows us.

The next logical step appears to take the piecemeal development to its logical end. I predict that in the future, analytics services will be increasingly partitioned and outsourced. The next research frontier is how an outsourcing process striving to totally decouple the stages of an analytics project should be managed.

REFERENCES

Articles

Apte, C.V., Hong, S.J., Natarajan, R., Pednault, E.P.D., Tipu, F., Weiss, S.M., (2003) "Data-intensive analytics for predictive modeling", BM Journal of Research and Development, Januray 2003, vol. 47, issue 1, pp. 17-23.

Baars, H., Kemper, H.-G. (2008) “Management support with structured and unstructured data – an integrated business intelligence framework”, Information Systems Management 25, pp. 132-148

Barton, D., Court, D. (2012) "Making Advanced Analytics Work For You", Harvard Business Review, October 2012, pp. 79-83.

Bayrak, T. (2015) “A Review of Business Analytics: A Business Enabler or Another Passing Fad”, Social and Behavioral Sciences 195 (2015), pp. 230-239.

Bose, R. (2009) "Advanced analytics: opportunities and challenges", Industrial Management & Data Systems, Vol. 109 Issue: 2, pp.155-172, https://doi.org/10.1108/02635570910930073

Chaudhuri, S., Dayal, U., Narasaya, V. “An overview of business intelligence technology”, Communications of the ACM, August 2011, Vol. 54, No. 8, pp. 88-98.

Chen, H., Chiang, R.H.L., Storey, V.C. “Business intelligence and analytics: From big data to big impact”, MIS Quarterly Vol. 36 No. 4, pp. 1165-1188, December 2012.

Davenport, T.H., Harris, J.G., De Long, D.W., Jacobson, A.L. (2001) “Data to Knowledge to Results Building an Analytic Capability”, California Management Review, Vol. 43, No. 2, Winter 2001.

Davenport, T.H. (2006) "Competing on analytics", Harvard Business Review, January 2006.

Delen, D., Demirkan, H. (2013), "Data, information and analytics as services", Decision Support Systems 55 (2013) pp. 359-363.

Dubois, A., Gadde, L.E. (2002) “Systematic combining: and abductive approach to case research”, Journal of Business Research 55, (2002) pp. 553-570.

Eisenhardt, K.M. (1989), “Building Theories from Case Study Research”, The Academy of Management Review, Vol. 14, No. 4, October 1989, pp. 532-550.

Evans, J.R. (2012), Business Analytics: The Next Frontier for Decision Sciences", Decision Line, Vol. 43, No. 2, March 2012.

Gandomi, A., Haider, M. (2015), “Beyond the hype: Big data concepts, methods, and analytics”, International Journal of Information Management 35 (2015) pp.137-144.

Holsapple, C., Lee-Post, A., Pakath, R. (2014), “A unified foundation for business analytics”, Decision support systems 64 (2014) pp. 130-141.

Kohavi, R., Rothleder, N.J., Simoudis, E. (2002), "Emerging trends in business analytics", Communications of the ACM, August 2002, vol. 45. No. 8.

Kowalczyk, M., Buxmann, P. (2015), “An Ambidextrous perspective on business intelligence and analytics support in decision processes: Insights from a multiple case study”, Decision Support Systems 80 (2015), pp. 1-13.

Lavalle, S., Lesser, E., Shockley, R., Hopkins, M.S., Kruschwitz, N. (2011) “Big Data, Analytics and the Path From Insights to Value”, MIT Sloan Management Review, Winter 2011, pp. 21-31.

Lee, A.S., A “Scientific Methodology for MIS Case Studies”, MIS Quarterly, March 1989, pp. 33-50.

Mortenson, M. J., Doherty, N.F., Stewart, R. (2015), “Operational research from Taylorism to Terabytes: A research agenda for the analytics age”, European Jounal of Operational Research 421 (2015) pp. 583-595.

Power, D.J. (2016) “Data Science: supporting decision-making”, Journal of Decision Systems, 2016, Vol. 25, No. 4, pp. 345-356.

Ranyard, J.C., Fildes, R., Hu, T-I. (2015), “Reassessing the scope of OR practice: The influences of problem structuring methods and the analytics movement”, European Journal of Operational Research 245 (2015) pp. 1-13.

Ranbotsham, S., Kiron, D., Prentice, P.K. (2015) "Minding the Analytics Gap", MITSloan Management Review, Spring 2015, pp. 62-68.

Seddon, P., Constantinidis, D., Tamm, T., Dod, H. (2017) “How does business analytics contribute to business value”, Information Systems Journal 27, pp. 237–269.

Shanks, G., Bekmamedova, N. (2014) “Achieving benefits with business analytics systems: an evolutionary process perspective”, Journal of Decision Systems, 21:3, pp. 231-244.

Shields, B. (2017) “Integrating Analytics in Your Organization: Lessons From the Sports Industry”, MIT Sloan Management Review, Winter 2018, vol. 59, issue 2, pp. 108-115.

Simon, H.A. (1955) “A Behavioral Model of Rational Choice” The Quarterly Journal of Economics, Volume 69, Issue 1, 1 February 1955, pp. 99–118

Tversky, A., Kahnemann, D. (1974) “Judgment under Uncertainty: Heuristics and Biases”, Science, vol. 185, 27 September 1974. pp. 1124-1131.

Viaene, S., Van den Bunder, A. (2011) “The Secrets of Managing Business Analytics Projects”, MIT Sloan Management Review, Fall 2011, Vol. 53, No. 1, pp. 65-69.

Vidgen, R., Shaw, S., Grant, D.B. (2017), Management challenges in creating value from business analytics", European Journal of Operational Research 261 (2017), pp. 626-639.

Walker, G., Weber, D. (1984) "A Transaction Cost Approach to Make-or-Buy Decisions", Administrative Science Quarterly, Vol. 29, No. 3 (Sep., 1984), pp. 373-391.

Books

Davenport, T.H., Harris, J. (2007), “Competing on Analytics”, Harvard Business School Press, Boston, Massachusetts.

Metters, R. D., King-Metters, K. H., Pullman, M., Walton, S. (2012) “Successful Service Operations Management”, Second Edition, Cengage, ISBN 978-81-315-1773-4.

Internet sources

Gartner 2018: https://www.gartner.com/it-glossary/advanced-analytics/ (visited on 27 February 2018)

INFORMS 2017: https://www.informs.org/Explore/Operations-Research-Analytics, visited on 23 December 2017)

Relex 2018: https://www.relexsolutions.com/stockmann-case-fi/ Visited on 23 February 2018

Sellforte 2018: Https://www.sellforte.com/pdf/Sellforte_Stockmann_Delicatessen_Promotion_Effectiveness_Case_Study_2017.pdf Visited on 23 February 2018

Appendix A: List of interview questions (in Finnish)

Tämän haastattelun tarkoituksena on kartoittaa edistyksellisen analytiikan implementaation prosessivaiheita, sekä analytiikan implementoinnin menestykseen vaikuttavia tekijöitä.

1. Mitä analytiikka ja edistyksellinen analytiikka tarkoittavat sinulle?
2. Onko analytiikan tyyppejä perusteltua erotella toisistaan?
3. Onko analytiikkahankkeita mahdollista tarkastella prosesseina, joissa vaiheet seuraavat toisiaan?
4. Mitkä ovat analytiikkahankkeiden vaiheet?
5. Mitä ajatuksia sinussa herättää analytiikkahankkeiden tarkastelu prosessina, jonka vaiheet ovat 1) datan keruu ja valmisteleminen analyysia varten 2) insight generation ts. datan analysointi 3) uuden tiedon tuottaminen ja esittäminen 4) johdon päätöksenteko 5) päätöksen implementaatio käytäntöön?
6. Mitkä tekijät edesauttavat analytiikkaprojektin onnistumista vaiheesta 2 eteenpäin?
    a. Insight generation ts. datan analysointi
    b. Uuden tiedon tuottaminen ja es ittäminen
    c. Päätöksenteko
    d. Päätöksen implementaatio käytäntöön
7. Mitkä ovat relevantteja mittareita yllä olevien vaiheiden onnistumisen arvioimiseksi?
8. Mikä on ihmisten vaikutus analytiikkaprojektin onnistumiseen?
    a. Pystytkö arvioimaan ihmisten vaikutusta organisaation eri tasoilla vaihe vaiheelta
        i. Johto
        ii. Työntekijät
        iii. Analyytikot (tiedon analysoija)
        iv. Analyytiikka-manageri (projektin johtaja, tulosten kommunikoija)
9. Kerro omin sanoin analytiikkahankkeiden menestykseen vaikuttavista tekijöistä